

ベイズ計量経済分析における最近の発展*

和 合 肇†

ベイズ手法は理論的な計量経済分析や統計分析において広く用いられているが、応用的な実証研究においては、まだそれほど広くは受け入れられていないのが現状である。これは、第1にベイズ計量経済分析には尤度関数と事前分布が必要である、第2に主観的である事前分布の定式化の問題、第3に多次元の積分計算の結果ほとんどの事後モーメントは求めるのが困難な場合が多い、というような理由のためである。このように応用研究の少なさは、実際に計算を行うには難しすぎるという認識によるものである。しかしながら、最近の物理的な計算力の急速な発達が発算能力の理論的な発展と相俟って、ベイズ手法が適用可能な問題の幅を急激に広げている。そこで、ベイズ的な方法の基礎について簡単に紹介した後、ベイズ手法を計量経済分析に応用したいくつかの研究について述べる。後半では、ベイズ計量経済分析における最近の数値計算面での発展を要約する。特に、最近多くの実証分析で使われるようになったマルコフ連鎖モンテカルロ法を用いた事後分布をシミュレートする方法について述べる。目的は、ベイズ手法は実際に計算可能であり、実証研究においても簡単に利用できることを経済分析に携わっている研究者に知ってもらうことである。

1. はじめに

計量経済学は、データを用いて経済的な現象についての確信（経済理論）を改訂する分野の学問である。ベイズ流計量経済学では、観測された事実をもとに確率法則にしたがってこの改訂が行われる。ベイズ流計量経済学で述べていることは、経済理論における主要な問題である期待効用最大化とベイズ学習過程と同じであるといえる。計量経済学が究極的に表明する問題は、経済モデルの中で経済主体が直面しているものと同じである。すなわち、新しく観測データが与えられると、次にどのような決定をすべきであろうか？ 2つの対立モデルを生成するデータを用いると、その相対的な妥当性はどのように変化するのか？

ベイズ統計学で採用されている手法は、理論的な計量経済分析や統計分析においてますます多くの注目を集めるようになってきているが、ベイズ手法を用いた実証分析に対する関心は、理論における関心の高さに比べて釣り合った増加を示していない。ごく最近までベイズ計量経済学の応用は、主として理論面への貢献に関心がある人々によって行われており、真にベイズ的である応用研究の部分は少ない。この理由はいくつかある。第1に、ベイズ計量経済分析は尤度関数と事前分布が必要である。他方、非ベイズ分析では必要ない。第2に、主観的な事前分布の正しさを弁護しなければならず、もし同意が得られなければその研究は無視されるかも知れない。第3に、ほとんどの事後モーメントは、必要な積分が評価できないために、求めるのが困難な場合が多い。このように応用研究が少ないということは、最も単純な場合を除いて、ベイズ手法は実際に計算を行うには難しすぎるという認識によるものである。しかしながら、計算力の最近の急速な発達は、計算能力の発達を可能にした理論的な発展と相俟って、ベイズ手

* 本稿は第61回日本統計学会チュートリアルセミナーの講演内容をもとに加筆したものである。その後、日本統計学会のベイズ統計学の応用に関する研究部会、東北大学経済学部、早稲田大学政経学部、大阪大学経済学部をはじめ、いくつかのセミナーでの報告に際しての参加者によるコメントに感謝したい。

† 新潟大学経済学部 〒950-2181 新潟市五十嵐2の町8050 (e-mail: wago@econ.niigata-u.ac.jp)

法が適用可能な問題の幅を急激に広げている。ベイズ的な計算手法を用いている多くの論文は、統計的な性質や特別な数学的な知識を利用しているので、応用計量経済分析を行っている人々にはまだあまり馴染みがない。

本論の目的は実際の面にある。ベイズ的な方法を採用する理由については、ベイズ的な方法と古典的な方法を比較して議論した優れたサーベイ(たとえば, Poirier, 1988, 1995; Berger, 1985 など)もあり, あまり深くは立ち入らない。むしろ実証研究にベイズ手法を適用することに集中する。全体の目的は, ほとんどあるいは特別な知識を持っていない研究者に, 応用実証研究におけるベイズ手法の有用性を認識してもらうことにある。

本論の構成は次の通りである。第2章で, ベイズ計量経済分析について簡単にサーベイする。第3章では, ベイズ的な手法を実証研究に应用到することに焦点を合わせ, ベイズ実証分析の優れた点(3.1節)と問題点(3.2節)について述べ, 3.3節はベイズ応用実証分析の例として, 単位根モデルとEBA(極値境界分析)を中心としてサーベイを行う。第4章以降では, ベイズ流の実証研究を行うのに用いられているいくつかの計算方法を中心として述べる。まずベイズ計量経済分析で用いる基本的な計算方法である正規近似(4.1節), 数値積分(4.2節), ラプラス近似(4.3節)について, その概略を説明した後, 第5章では, 最近多く用いられるようになったモンテカルロ積分について, 単純な場合(5.1節)とインポートランス関数を用いた場合(5.2節)について, その使用例とともに解説する。Poly- t 密度について簡単に触れた後(5.3節), 第6章ではもっと複雑な問題に用いることができるマルコフ連鎖モンテカルロ・シミュレーションについて述べる。一般的ベイズ推定(6.1節)とマルコフ連鎖(6.2節)についてふれた後, 代表的な方法であるMetropolis-Hastingアルゴリズム(6.3節)とよく使用されるGibbsサンプリング法(6.4節), それにデータ拡大アルゴリズム(6.5節)について述べる。次に, 第7章ではMCMCシミュレーションを計量経済分析にどのように適用するかを応用例とともに述べる。ここでは, 事後分布からシミュレーションを行う方法の計量経済分析への応用について強調し, 重要な問題である収束の診断方法や数学的詳細は省略する。最後に, 第8章ではいくつかの結論を述べ, 付録でベイズ実証分析を行う際に役立つコンピュータ・パッケージについて述べる。

2. ベイズ計量経済分析

最初に, ベイズ計量経済理論の背後にある基本的な理論的概念について簡単に整理しておく。最近では, 英語で書かれた多くの計量経済学のテキストではいくつかの章でベイズの方法が扱われているが, 残念ながら日本語で書かれた計量経済学の教科書ではほとんどふれられていないのが現状である。ベイズ計量経済理論についてより詳しく知りたい場合は, このトピックスについての多くの参考文献があるので, それらを参考にすると良い(たとえば, Zellner, 1971; Poirier, 1988, 1995)。

ベイズ計量経済学は, 次のベイズの法則に基づいている。

$$p(\theta|y) = \frac{p(\theta)L(y|\theta)}{p(y)} \propto p(\theta)L(y|\theta)$$

ここで θ はパラメータベクトル, y はデータベクトルあるいは行列, $p(\cdot)$ は確率密度関数を表す。 $p(\theta)$ はデータを分析する前にパラメータについての主観的な確信を表す**事前密度**(prior pdf), $L(y|\theta)$ は θ に特定の値が与えられたときの y の条件付き密度で**尤度関数**(likelihood function)と呼ばれる。そして $p(\theta|y)$ は θ の**事後密度**(posterior pdf)であり, データに含まれる新しい情報を考慮して事前の確信を更新することを表している。事後密度は, 観測された

データ y の条件付で θ について得られるすべての情報を要約したもので、 θ に関する推測は事後密度から得られる。たとえば、 θ の事後平均 $E(\theta|y) = \int \theta p(\theta|y) d\theta$ は、 θ の点推定値としてしばしば用いられる。

予測も簡単に得ることができ、 y^* を 1 組の将来の観測値とすると、予測は予測密度 $p(y^*|y)$ をもとにして行われる。確率法則を用いて、 $p(y^*|y) = \int p(y^*|\theta, y) p(\theta|y) d\theta$ となり、ここで $p(y^*|\theta, y)$ は将来観測値に関する尤度関数、そして $p(\theta|y)$ は事後密度である。

ベイズの方法を用いて仮説を比較するいくつかの方法がある。その中で最もよく使われるのは事後オッズ比 (posterior odds ratio) である。たとえば、もし 2 つの異なる仮説 H_1 と H_2 (入れ子型あるいは非入れ子型) を比較しようとするれば、事後オッズ比 K_{12} は $K_{12} = p(H_1|y)/p(H_2|y)$ となる。言い換えると、 K_{12} は観測データから得られる 2 つの仮説の確率の比である。もし θ_1 と θ_2 が、 H_1 と H_2 それぞれの仮説のもとで妥当なパラメータベクトルであるなら、事後オッズ比は次のように計算される。

$$K_{12} = \frac{\pi(H_1) \int p(\theta_1|H_1) L(\theta_1|y, H_1) d\theta_1}{\pi(H_2) \int p(\theta_2|H_2) L(\theta_2|y, H_2) d\theta_2}$$

K_{12} の計算には、仮説が成り立つ事前確率 ($p(H_i), i=1, 2$) と同様に、各仮説のもとでの事前情報と尤度が必要である。1 つの決定は、 $K_{12} > 1$ なら H_1 を受容し、その他の場合は H_2 を受容することである。もちろん、もっと手の込んだ方法も利用できる。たとえば、非線形制約の検定に関心がある場合、制約のないモデルと制約付のモデルの両方の下で事前分布を作成するのは大変であるので実際的ではない。しかし、McCulloch and Rossi (1992) は、この種の問題を克服できるある優れた方法を開発している。

応用研究ではベイズファクターも通常報告される。これは 2 つの仮説に対する尤度比のサポートになる。言い換えると、ベイズファクターは 2 つの競合する仮説に対して同じ事前分布を与えて計算した事後オッズ比である。

計量経済学で通常行われる代表的な統計的分析であるパラメータの推定、仮説検定、そして予測に対するベイズ的な方法を簡単に要約する。ベイズ的な方法と従来からの古典的な計量経済分析の方法との重要な違いは、前者が未知の変量 (たとえば θ, y^* や仮説の真実性など) を確率変数として扱い、既知の変量 (たとえば、データ) に条件づけるのに対し、後者はデータを確率変数として扱い、パラメータや仮説の状態を非確率的なものとして考えることにある。ベイズ分析を行うときには、情報のない事前分布が用いられる場合が多いけれども、尤度関数と事前分布の両方を定式化する必要がある¹⁾。定式化を行った後、パラメータの推定値を得たり、仮説を検定したり、あるいは予測を行ったりするために、簡単に上で述べた方法を使うことができる。したがって、概念的には少なくともベイズ的な方法は、古典的な方法に比べればずっと簡単に使うことができる。

ベイズ法を用いる場合には、積分の計算が必要になる。事後平均はしばしば推定量になり、 $E(\theta|y) = \int \theta p(\theta|y) d\theta$ を計算する。事後オッズを得るためには、 $\int p(\theta|H_1) L(\theta|y, H_1) d\theta$ を計算する必要がある。自然共役事前分布²⁾を持つ正規線形回帰モデルのような標準的な場合では、これらの式の解析的な解は存在する。しかし、一般的な場合には存在せず、ベイズ分析を行う場合はコンピュータ集約的な手法に頼らざるを得ない。

¹⁾ ベイズ意思決定理論的な分析では、損失関数も定式化する必要があるが、この段階は実際にはよく省略される。

²⁾ 事前密度は、事後密度と事前密度が同じクラスの確率密度関数になるようにしばしば選ばれる。このような事前密度は、自然共役事前分布と呼ばれる。

2.1 事前分布

事前分布は、分析者が関心をもつパラメータ θ についての確信の度合いを記述したもので、現在のデータセットからすべての情報が得られる以前の情報である。事前分布はそれが確信の度合いを表すことから、厳密な意味では不正確ではあり得ないが、確信が合理的であることの検定が必要であるという意味で主観的である。しかし、事前分布はデータセットからのどんな情報も拒否するほど独善的であるべきではない。

事前分布は次のようないくつかの理由によって分類される。最も簡単な事前分布を作成する方法は、とりうるパラメータの値を集めて全体が1になるように基準化することによって、事前分布をできるだけ事前の確信を反映させるようにする。 θ が離散的な値をとる場合は自然な方法であり、連続的な値をとる場合は実線上の区間をとって θ に対する事前のヒストグラムを作成する。あるいは θ に対する事前分布が、あるパラメトリックな分布族 $\pi(\theta|y)$ に属すると仮定して、 y を結果ができるだけ真の事前の確信に近くなるように選ぶ。たとえば、 y が2次元の場合、2つのモーメント（たとえば平均と分散）か、2つの分位値（たとえば50%と95%）を与えれば、その厳密な値を決めることができる。しかしこの方法の限界は、事前の確信が標準的なパラメトリック型に収まらない場合がある点である。さらに、一見すると同じように見える2つの分布が、全く異なる性質を持っている場合がある。たとえば、Berger (1985, p. 79) で指摘されているように、Cauchy (0, 1) と Normal (0, 2.19) 分布は同じ25%, 50%, 75%点 (-1, 0, 1) を持ち、プロットすると非常に似たように見える密度関数が全く異なる事後分布になることが知られている。

共役事前分布 (conjugate prior) は、尤度関数（とそのパラメトリック型）と結合したとき解析的に調べることができる形（すなわち、多変量 t 分布のような既知の認識できる形）を持つ事後分布になる。分布を解析し、積分を評価する数値的な方法における最近の理論的發展がある前は、共役事前分布はベイズ分析で非常に重要であった。その事前分布が仮定した尤度と共役な特定の分布のクラスの制約の中で、よく当てはまるようにする必要があった。事後分布を形成する数値手法に新しい發展があった後は、共役事前分布は必要性がなくなり、自由に自らの事前の確信を真に表しているどんな事前分布も選べるようになった。事実、最近ではリーズナブルと思われる自由な事前分布を描けるようになり、コンピュータを使って様々な分布（正規分布、 t 分布、ベータ分布等）を描き、分布の形をコントロールするパラメータを適当に変えることによって、自分の確信と合うように分布を作成することができる。これは事前情報の数学的表現を与え、少なくとも共同研究者の1人が様々な統計分布を描くことができれば、かなり容易に行うことができる。

事前分布は、少なくとも θ のいくつかの要素に関して事前に持っている知識を表す**報知的 (informative)** であるか、あるいは無知の状態を表す**非報知的 (non-informative)** に分けられる。最もよく知られた非報知事前分布、すなわち無知である事前分布は **Jeffreys 型事前分布** で、情報行列（すなわち、パラメータ θ に関して対数尤度関数の2階の偏導関数の負の期待値）の行列式の平方根に比例してとられる。誤差項が正規分布にしたがう古典的な線形回帰モデルでは、 $\theta = (\beta', \sigma^2)$ であり、Jeffreys 型事前分布は $h(\theta) \propto \sigma^{-1}$ で与えられる。簡単に言えば、Jeffreys は、もし事前の無知を表したければ、無限の許容可能な範囲を持つパラメータ（すなわち、一様分布）に対しては定数に比例するような事前分布を使うべきであり、 $[0, \infty)$ の許容できる範囲を持つ（標準偏差のような）パラメータに対してはパラメータの自然対数上で一様分布である事前分布を用いるべきであることを提案した。古典的な回帰モデルに対する事前分布につながるのはいずれの法則である。

事前分布は**正則 (proper)** か、あるいは**非正則 (improper)** である。正則な事前分布は、 θ の

許容可能な範囲に関して積分すると1になる。報知的である分布は、適切に基準化すればほとんど常に正則である。たとえば、もし切断された正規分布を限定された範囲のパラメータに関する事前分布として用いると、この分布は切断を考慮して基準化しなければならず、事前分布全体が1になるように分布を基準化し直す必要がある。非報知事前分布は、しばしば非正則である。無限の範囲の一様分布は積分すると無限になり、1にはならないので非正則である。非正則事前分布は仮説検定分野で問題が生じる。これは、事後オッズ比を任意に基準化することができるために、競合モデルに対する相対的な支持の評価が不可能になるためである。このために、非正則事前分布をできるだけ使用しないようにする。多くの応用分析では、データの尤度関数によって事後分布をほとんど完全に決めることができ、他方、仮説検定に対してはオッズ比が任意に基準化されないような弱い報知事前分布を作ることができる。

2.2 回帰モデルの標準的な分布

計量経済分析では線形回帰モデルがよく使われるが、通常用いられている散漫な事前分布の下で、ベイズ分析で用いるいくつかの標準的な分布を示すことにする。必ずしもそのような非報知事前分布を常に用いる必要はないが、ベンチマークとしては意味があり、回帰パラメータの事後分布を解析的に調べるのに役立つ。さらに、本節で示される分布は、4章以降で述べるいくつかの数値積分法で確率抽出する乱数に関して役に立つ情報を与える。

まず、すべての古典的な仮定の下での標準回帰モデルを考える。

$$y = X\beta + \epsilon$$

ここで y は $(n \times 1)$ 従属変数の観測値ベクトルでその変動をモデルで説明しようとし、 X は $(n \times k)$ 説明変数行列、 β は $(k \times 1)$ の推定されるベクトルすなわち回帰パラメータ、そして ϵ は分散が σ^2 のホワイトノイズ誤差項の同じ次元のベクトルである。もし、回帰パラメータに非正則な一様事前分布(これも *Jeffreys* 型事前分布)、そして σ に標準的な *Jeffreys* 型事前分布を選択すると、 (β, σ) に対する同時事前分布は次のように書ける。

$$p(\beta, \sigma) \propto \frac{1}{\sigma}$$

正規性の仮定の下で、従属変数の観測値に対する尤度関数は

$$p(y|\beta, \sigma, X) = (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2}(y - X\beta)'(y - X\beta)/\sigma^2\right]$$

であり、同時事後分布は次のように表せる。

$$\begin{aligned} p(\beta, \sigma|y, X) &\propto (2\pi)^{-n/2} \sigma^{-(n+1)} \exp\left[-\frac{1}{2}(y - X\beta)'(y - X\beta)/\sigma^2\right] \\ &\propto (2\pi)^{-n/2} \sigma^{-(n+1)} \exp\left[-\frac{1}{2}(n-k)s^2/\sigma^2 + (\beta - \hat{\beta})'\sigma^{-2}X'X(\beta - \hat{\beta})\right] \end{aligned}$$

ここで、上の式は二次形式の項に分けられ、 β と σ^2 の最小2乗推定量 $\hat{\beta}$ と s^2 を用いて表されている。パラメータの周辺事後分布は

$$p(\beta|y, X) \propto [(n-k)s^2 + (\beta - \hat{\beta})'X'X(\beta - \hat{\beta})]^{n/2}$$

となり、これは多変量 t 分布にしたがう。そして

$$p(\sigma|y, X) \propto \sigma^{-(n-k+1)} \exp\left[-\frac{1}{2}(n-k)s^2/\sigma^2\right]$$

は逆ガンマ分布にしたがい、 σ^2 の周辺事後分布は χ^2 分布を基準化したものであることを意味している。分散-共分散行列を Σ と表す一般化誤差分布の場合は、この分布は簡単に一般化でき、 Σ の周辺事後分布は逆ウィッシュャート分布にしたがう。この分布は、逆ガンマ分布を一般化したものである。

3. ベイズ的な方法の実証分析への応用

3.1 ベイズ的な方法の利点

本稿での主要な関心は、ベイズ的な方法を経済問題に対する実証研究に応用することである。ベイズ的な見方と古典的な見方との違いをはっきりさせるために、両者の考え方と方法論に主に焦点を合わせた分析はPoirier (1988)あるいはBerger (1985), 118-126ページを参照すると良い。実証的な問題の分析に関しては、ベイズ的な計量経済分析の方法は古典的な方法と比べて、その単純性 (*simplicity*) と整合性 (*coherency*) という点において優れている。古典的な手法に必要とされる多くのテクニックと基準とは異なり、ベイズ的な方法では(非報知的な)事前密度と尤度関数だけを必要とする。対照的に、古典的な実証分析では適切な手法を選択し、その統計的な性質を知らなければならない。しかも多くの性質は漸近的にのみ既知であり、小標本でのパフォーマンスにともなう問題は残されている。

ベイズ法は観測されたデータの条件付の事後密度に基づいている。これは、古典的な手法が推定量の性質が繰り返し標本で評価される不偏性のような基準に基づいているのと対照的である。このような基準はベイジアンにとってはほとんど関心がない。さらに、この条件付けを行うことは、特定の応用に対して合理的な計量経済理論を単純化するだけでなく、しばしば応用そのものを単純化する。たとえば、Geweke (1988)の不等号制約にしたがうパラメータをもつモデルのベイズ分析は、古典的な方法に比べて非常に簡単になるのはよい例である(Poirier, 1988)。同様に、McCulloch and Rossi (1992)は、ランダム係数構造をもつ多対プロビットモデルを用いて計算的に可能な方法を開発した。パネルデータを用いた研究では、消費者の異質性を認めることが必要であり、そのような異質性をモデル化する標準的な方法は、ランダム係数モデルである。しかし、古典的な方法を用いて多対プロビットモデルの枠組みでランダム係数モデルを用いると、多くの計算上の厄介な問題が生じる。しかしGibbsサンプリングを用いたベイズ法では、階層モデルに対してと同様な利点がある。

ベイズの方法による条件付けを行った場合の性質は、観測された標本に対して厳密な結果が得られることを意味している。しかし、古典的な計量経済学では、特定の検定や推定量を用いるのに小標本での性質が漸近的な結果に対応していると考えて、しばしば漸近的な性質を用いて正当化する。小標本と大標本の性質が大きく異なる場合がたくさんあるが、推定量の小標本の性質を得るのは難しい場合が多い。

ベイズ法のもう1つの利点は、局外母数 (*nuisance parameter*) の取り扱いにある。これは古典的な計量経済分析を行うには特に厄介な問題である。通常は、それらを有限標本の性質を失うというコストを払って一致推定量で置き換える必要がある。しかしベイズ法では局外母数に関して簡単に積分できる。たとえば、線形回帰モデルで回帰係数 β と誤差分散 σ^2 とすると、事前分布に尤度関数を掛けて事後分布 $p(\beta, \sigma^2|y)$ を得る。もし β だけが必要なら、ベイズ分析では $p(\beta|y) = \int p(\beta, \sigma^2|y) d\sigma^2$ とするだけである⁹⁾。

3.2 ベイズ的な方法の問題点

事前分布の問題

ベイズ法で一般に認識されている問題点は、事前情報を含めることを認めることである。この事前情報を用いることに対して多くの論争が行われたが、その詳細についてここでサーベイを行わない。しかし、もし事前情報が得られればそれを含めるべきであり、ベイズ法はそれを行う最良の方法を提供する。

実証分析では、事前分布を変化させたことによる事後分布の感度問題は重要である。しかし、事前分布の生成 (*elicitation*) の問題はより重要である。計量経済分析では自分だけが満足する結果が得られればよいわけではない。このために、自分の事前の確信に対応するただ1つだけの事前分布を選択すべきでなく、むしろ、様々な専門家の意見を取り込んだ一定の範囲の事前分布を選択すべきである。ほとんどの応用問題では、それまでの研究や経済理論から研究者間にかかなりの合意がある。たとえば、 β を生産モデルでの規模に対する収穫を測るパラメータとすると、研究者によって様々な事前分布を β に持つ。ほとんどは β が 1 に近く、人によっては β が 0.75 に近いと思っているが、だれも β が 100 に近いと思っていない。そこで、合理的な事前密度は主要な確率が区間 $[0.5, 1.5]$ 内にあるとする。しかし、 $N(1, 0.25)$ 、 $N(0.75, 1)$ 、あるいは $N(1.25, 1)$ のように、この性質を持つ多くの事前分布があるので、応用研究ではこれらすべての様々な事前分布を試すことができる。尤もらしい一定の範囲の事前分布に対して結果が頑健であることは、ただ1つの事前分布に対して示される結果よりもずっと受け入れやすい。多くの実証分析では、推定は事前分布の選択に関してそれほど敏感でない。もし異なる合理的な事前分布によってまったく異なる推定値が得られたとすると、重要な情報が尤度関数の性質について明らかになる。すなわち、尤度関数は典型的にはかなりフラットで、したがって非報知的であることになる。

もし事前分布の生成が期待できないか非常に困難である場合には、非報知事前情報のクラスをそのかわりに用いることができる。

多くの場合、パラメータについてたぶん曖昧な事前情報しか持っていない場合が多い。再び、 β を規模に関する収穫を測るパラメータとすると、 β は 1 に近いと感じており、1.5 より大きいか 0.5 より小さくはないというはっきりとした確信を持っている。この性質を持つ非常に多い事前分布がある (たとえば $N(1, .0625)$)。実際には、多くのベイズ的な応用研究では扱いやすい事前密度の関数形が選択される。次に、事前情報に基づいて事前密度のパラメータ (ハイパーパラメータと呼ばれる) を選ぶ。したがって、以前の例での β に対する事前分布は正規分布であると決める。それはたんに正規分布が扱いやすいからである。そして、事前の情報に基づいて事前平均が 1 で事前分散が 0.0625 とする。 β と σ^2 によって標準的な方法で定式化された正規線形回帰モデルに対しては、 β と σ^2 に対する正規・ガンマ事前分布は自然共役になり、解析的な結果が存在し、この事前分布がこのモデルに対して広く用いられるようになった。

³⁾ 実際には、局外母数に関して積分すると問題をずっと簡単にできる。しかし、次の2つの理由から結果に影響を与えることがある。それは i) 古典的な計量経済分析では、局外母数を一致推定値でしばしば置き換える方法があるが、小標本では一致推定量によって結果が異なることがあり得る。しかし、そのような類似の問題はベイズアンでは生じない。ii) 局外母数のベイズアンでの扱いは、周辺密度 ($p(\beta|y)$) で考えることになる。古典的な手法に似たベイズ的な方法では条件付密度 ($p(\beta|\sigma^2=s^2)$ 、ここで、 s^2 は σ^2 の一致推定値) で考えることである。実際には、古典的計量経済学では滅多に $p(\beta|\sigma^2=s^2)$ を使わない。いうまでもなく、局外母数に一致推定値を当てることは小標本効果を持つ。

⁴⁾ 事前分布の感度分析についてのベイズアン文献は急速に増加している。この分野への役に立つ入門書として Berger (1985), 195-247 ページ、あるいは Kadane (1984) がある。線形回帰モデルに対する事前分布の感度についての議論は、Leamer (1978, 1983, 1985) にある。

ある関数形が事前分布に対して選ばれたら、事前分布の導出問題は事前のハイパーパラメータの選択に要約される。実際は、ベイジアンは、パラメータの経済学的な解釈を通じて非公式に推論することによって、事前のハイパーパラメータをしばしば選択する。たとえば、上の例で β は規模に対する収穫の弾力性であるから、それはたぶん 1 に近いと推論する。したがって、1 の事前平均と 0.0625 の事前分散を選択する。モデルのパラメータに単純な経済学的解釈ができない場合が数多く存在するが、合理的な事前分布を導出することはそれでも可能である。たとえば、Koop and Poirier (1993) はロジットモデルに対する自然共役事前分布を導出している。彼らは最初に、このモデルのパラメータを直接に解釈するのが難しく、したがって事前分布の導出が難しいと述べている。それでも、それらの事前分布に対するハイパーパラメータを、パラメータ自体ではなく観測可能性に関して推論することによって、選択する方法を述べている。すなわち、事前分布の導出を次のような問いをもとに行う：「ある特性をもつ個人は、その特性をもたない人に関連する別の特性を選択することがどのくらいあり得るか？」

Kadane et al. (1980) は、正規線形モデルに対してこの種の接近方法をとるソフトウェアを開発している。この対話型ソフトウェアは使用者に、 X のある値に対して y のとる値を予測するように促す。すると想定した事前分布が、いくつかの質問に対する反応を基に計算できる。

これまででは、事前分布の選択に関するいくつかの問題に対する簡単な紹介を行った。その他の事前分布の作成方法は Berger (1985) の第 3 章にある。この段階でいくつかの点を強調しておく必要がある。i) 多くのモデルのクラスでは、通常用いられる事前分布の標準的なファミリーが存在する。たとえば正規線形回帰モデルに対する正規・ガンマ事前分布である。これらのクラスに対しては、事前分布の作成問題は適切な事前のハイパーパラメータを選択することに帰着する。ii) 多くの人の納得を得るために、分析者は事前分布の感度分析を行うべきである。iii) そのような事前情報を含めることは、合理的な非報知事前分布がほとんどのモデルに対して存在するので、どうしても必要なわけではない。

後者の場合、そのようなクラスの非報知事前分布は非常に役に立つ。ほとんどのモデルに対して、様々な定義の非報知性に対応する非報知事前分布があるが、実証的な結果は通常そのような事前分布の選択に対して敏感ではない。最もよく知られた非報知事前分布は、 $p(\theta) = \text{定数}$ である。そのようなフラットな事前分布を用いることは、尤度関数に比例した事後分布を得ることになり、したがって、これを用いるベイズ分析は尤度関数だけを含むことになる。このようなフラットな事前分布を用いた結果は、古典的な研究で得られたものとはしばしば似通った結果になる⁹⁾。たとえば、事後密度のモードは最尤推定値と同じである。

フラット事前分布は、パラメータ変換に対して不変ではないことで批判されている。すなわち、 θ がある区間に入る事前確率は、 $f(\theta)$ を任意の関数としたときに $f(\theta)$ がその区間に入る最大確率と同じでない。非報知事前分布を導出するもう 1 つの方法はジェフリーズのルールとして知られており、事前分布を情報行列の行列式の平方根に比例する様に選択することを示唆した。このようにして作成された非報知事前分布は不変であり、多くの場合結果として得られた事前分布はフラットなものであることに注意する。たとえば、正規線形回帰モデルではジェフリーズのルールの応用は、 β と $\ln(\sigma)$ に対するフラット事前分布と同じである $p(\beta, \sigma^2) \propto \sigma^{-1}$ の形の前分布になる⁹⁾。

⁹⁾ しかし、これは常に成り立つわけではない(たとえば、以下の単位根計量経済学についての議論を参照)。大雑把な経験則では、フラットな事前分布を用いたパラメータ推定値は最尤推定値と非常に似た数値になる傾向にあるが、仮説検定はしばしば非常に異なる結果を生み出す。この論文の第 1 節では、ベイジアン仮説検定に含まれる方法を簡単に述べている。概観した方法は古典的なものとは全く異なり、このためにかなり異なった結果になることを強調する価値がある。

しかし、別のもっと深刻な非報知事前密度の問題点は、それらがしばしば非正則になる、すなわち積分して1にならないことである。実線上での一様密度が通常用いられているが、非正則な非報知事前分布である。非正則な事前分布は推定や予測に対しては問題を起こさないが、同じことは仮説検定に対しては当てはまらない。特に、制約のない対立仮説に対するシャープな帰無仮説の検定に対する事後オッズは、しばしばデータに拘わらず帰無仮説が好まれる(Leamer (1978) 参照)⁷⁾。結果として、非正則な事前分布は、もし研究者が事後オッズ比で一点帰無仮説を検定したいならば、用いるべきではない⁸⁾。

3.3 ベイズ法を用いた応用分析のサーベイ

ベイズ流の分析は応用経済分析への研究には僅かしか利用されてこなかったけれども、それが受け入れられたのは90年代になってからである。初期の応用研究は、最近のいくつかの *Journal of Econometrics* のベイズ応用分析の特集号⁹⁾や、*Advances in Econometrics* に見られる¹⁰⁾。本稿ではベイズ分析を採用したすべての応用分野の詳細なサーベイは行わないが、計量経済分析についての包括的な情報は Zellner (1983, 1985) を参照すると良い。そのかわり、ここではベイズ法が広く用いられており、古典的な方法と比較可能な2つの分野である単位根検定とEBA(極値境界値分析)を取り上げることにする。前者は、多くの論争と有意な差がベイズ法と古典的な方法との間に見られるので、多少詳しく説明する。一方、後者はベイジアンと古典的な計量経済学者が多く的一致する立場に立てるケースの良い例である。このほかにも、制約付回帰分析の応用でベイズ的な方法が成功した例としてシラー・ラグがあるが、前に別の論文で取り上げたので省略する。

単位根問題

ある系列が階差定常過程(DSP)かトレンド定常過程(TSP)かという問題は、1980年代に多くの計量経済学者の注目を集めたトピックスであるが、経済学と統計学の両方の意味を持っている。系列が単位根を含むかどうかはマクロ経済学にとって非常に重要であり、技術パラメータが単位根にしたがうと考えると、リアル・ビジネスサイクルの理論の検証に多くの経済学者が参加した。統計的な問題としては、トレンド除去法と単位根検定の問題がある。古典的な方法による単位根問題の分析はすでに数多く行われており、それらの結果はHatanaka (1996)にまとめられている。

経済理論はいくつかの時系列は単位根を必ず含むはずであることを意味しているが、一方統

⁶⁾ 非報知事前分布についてのより詳細な議論については、Zellner (1971), 41-53 ページ、あるいは Berger (1985), 82-89 ページを参照。

⁷⁾ 事後オッズは、帰無仮説と対立仮説の下での尤度関数の加重平均の比であり、ここでウェイトは事前密度によって与えられる。もし対立仮説の下での事前密度がフラットになったら、尤度関数の裾にもっと大きなウェイトを与える。したがって、対立仮説の下で加重平均は小さくなり、帰無仮説は事後オッズ比では好まれるようになる。

⁸⁾ もっと正確には、事前密度は、帰無仮説の下では制約付であるこのようなパラメータに対しては非正則にならない。しかし非正則な事前密度は、帰無仮説と対立仮説の両方の下で制約のない、このようなパラメータに対しては問題を生じない。たとえば、線形回帰モデルでの関心が回帰係数 β がゼロかどうかを検定するのに事後オッズを用いるときに、 β_i に対しては正則な事前密度を使用しなければならない。しかし、 $\beta_j (j \neq i)$ に対してや誤差分散に対して、非正則な事前密度を使用することは完全に受け入れることができる。

⁹⁾ Bayesian Empirical Studies in Economics and Finance, ed. by D. J. Poirier, 1991; Bayesian and Classical Econometric Modelling of Time Series, ed. by L. Bauwens and M. Lubrano, 1995; Bayes, Bernoullis, and Basel, ed. by L. Bauwens, W. Polasek and H. van Dijk, 1996 の各特集号に多くのベイズ流の応用分析がある。

¹⁰⁾ T. B. Fomby and R. C. Hill ed. *Advances in Econometrics, Vol. 11 A, B* に多くのベイズ応用分析がまとめられており、Part A は Bayesian Computational Methods and Applications, Part B は Bayesian Methods Applied to Time Series Data のそれぞれ特集号である。

計理論は定常な系列と単位根系列の性質の違い、そしてそれに基づく推定量と検定統計量を強調している。Dickey and Fuller (1979) と Nelson and Plosser (1982) を始めその他のよく知られた研究以来、古典的な計量経済学者は経済時系列における単位根の検定問題にかなりの努力を注いできた。単位根計量経済学をめぐる多くの問題は、次の $AR(p)$ を再定式化した単純なモデルを考えることによって示すことができる。

$$y_t = \rho y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + u_t$$

ここで y_t ($t=1, \dots, T$) は観測されたベクトルの t 番目の要素であり、 u_t は観測されない誤差項である。実証が可能な単位根は $\rho=1$ ならば存在する。しかし、誤差項の性質について異なった仮定をすることが多い。たとえば、Dickey and Fuller (1979) は u_t が *iid.* 正規を仮定し、 t 分布に従わない $\rho=1$ を検定するのに、 t 統計量に対する臨界値をシミュレートした。その後 Phillips and Perron (1988) では、誤差項についてもっと一般的な仮定を考え、 t 統計量の関数である検定統計量を導いた。実証研究で単位根を検定するときには、しばしば Dickey-Fuller テストか Phillips-Perron の拡張が用いられる。しかし、いくつかのモンテカルロ研究からの事実 (たとえば Schwert, 1989 や DeJong, Nankervis, Savin and Whiteman, 1992 参照) は、これらの古典的な単位根検定は極めて低い検出力しかもたず、特にトレンド定常な対立仮説に対しては検出力は低いことを示している。

Sims (1988) と Sims and Uhlig (1991) による議論の発端となった論文は、単位根仮説を調べるのに古典的な計量経済学的方法を用いることを批判し、フラットな事前分布を用いることを支持しており、最近の多くの実証分析では、フラットな事前分布を用いたベイズ法を採用していることを示した。これらの論文は、ベイズ的な方法による単位根計量経済分析に大きな関心を引き起こした。最近の実証分析を見ると、フラットな事前分布を用いるベイズ法では、経済時系列の単位根仮説に関する推測が古典的な検定による結果とはしばしば異なることを示している。特に、DeJong and Whiteman (1991 a, b, c) によるフラットな事前分布を用いたベイズ分析では、Nelson and Plosser (1982) によって研究されたマクロ時系列や、Shiller (1981) と Marsh and Merton (1986) で分析されたストック価格や配当系列に単位根はほとんど見い出されていない。同様に、Schotman and van Dijk (1991a) によるフラットな事前分布を用いたベイズ分析では、実質為替レートや実質 GNP データに単位根があるという事実は、古典的な単位根検定の結果によって支持されよりもずっと弱いことを見い出した。

ベイズ単位根の分析は、ベイズ分析を行っている人々の内から批判されている。たとえば DeJong and Whiteman (1991a, b) で用いられたフラット事前分布の型は非報知ではないということについての論争から、Peter Phillips は時系列モデルにフラット事前分布を用いるのに強く反対した。それは、フラット事前分布はある種の不変性を欠いており、自己回帰係数がコログラムの形状に影響を及ぼし、データやその標本積率に含まれる期待情報量に影響を与えるという経路で、一般的なモデルに基づいた重要な情報を無視しているからである。Phillips (1991) は、非報知事前分布の作成に対する Jeffreys のルールを用いて、**無知の事前分布 (ignorance prior)** と呼ばれる事前分布を開発した。Phillips の事前分布は、拡散的な動きを意味するパラメータ空間に様々な事前ウェイトをおくという点で DeJong and Whiteman のものとは違う。この反直感的な性質は、Phillips の事前分布作成法の性質から生じるものであり、これは事前分布がモデルの標本理論的性質をまねていることを保証するものである。すなわち、もし ρ が実際に 1 より大きかったとすれば、モデルから生成されたデータはこの系列がすぐに発散する性質があることは明らかになる。したがって、一般的に 1 より大きい ρ の値はより多くの情

報を含んでおり、Phillipsの無知の事前分布が表している特性である。これらのモデルに基づいた事前分布を用いて、ベイズ法による事後解析は、単位根の有無について古典的な有意性検定と同じくらい不確実であることを示した。加えて、PhillipsはJeffreys型事前分布を利用した方法をNelson-Plosserデータに応用し、もっと一般的なJeffreys型の前分布を採用し、同様な結果を得ているDeJong and Whiteman (1989a)やZivot and Phillips (1991)で報告されたよりは、単位根モデルに対して決定的な証拠がないことを見出ししている。

Phillipsの無知の事前分布は、ベイジアンからも非ベイジアンからもともに批判された(Phillips, 1991に対する討論を参照)。この事前分布はパラメータ空間の非定常区域に非常に多くの事前ウェイトをおくので、非定常である事後確率はそれに応じて増加する。結果的として、多くのマクロ経済時系列に単位根があるという事実は、その他のベイズ流の分析で行った研究よりも多くなることをPhillipsは見出した¹³⁾。ある意味では、Phillipsの研究はベイジアンと古典的な方法との架け橋と考えることができる。これは公式的にはベイジアンであるが、その基礎にある考え方の多くは古典的な単位根とベイジアンの結果との乖離を説明することに捧げられている。時系列モデルで合理的な事前分布を導出するものは何かという点については、まだ回答が得られていない。非常に多くの研究が現在もこの分野で行われつつある。

DeJong and Whiteman (1991a, b), Phillips (1991) それにZivot and Phillips (1991) によって行われた単位根仮説のベイズ分析では、モデルのパラメータに関して線形である推定された定数項とトレンドをもつ共通のAR(k)モデルを用いている。Schotman and van Dijk (1991a)はPhillipsへの返答で、定数項とトレンドのあるAR(k)モデルの非線形の観測されない要素モデルを用いて、モデルのパラメータに関してフラット事前分布とJeffreys型事前分布の動きと、主要な自己回帰係数 ρ の結果として得られた事後分布を調べた。彼らは、Schmidt and Phillips (1989, 1992)におけるように、この要素モデルによる表現は単位根仮説の分析に対してより適切であるとした。それは要素モデルにおける確定項パラメータは、単位根帰無仮説とトレンド定常対立仮説の下では同じ解釈ができるが、線形モデルでのパラメータでは同じ解釈ができないからである。しかし、彼らの見方はDeJong and WhitemanやPhillipsによるものとはかなり異なる。特に、要素モデルの中ではフラット事前分布と比較した場合、Jeffreys型の前分布は単位根モデルのウェイトを小さくすることを示した。そして、DeJong and Whiteman (1991a, b)の結果と対照的に、フラット事前分布から導かれた事後分布の推定は確率1で単位根仮説になることも示した。さらに、単位根モデルを受け入れるという事後オッズ比は、レベルパラメータに非正則な事前分布を用いると発散することを示した。レベルパラメータに対するJeffreys型の前分布は非正則であるから、Schotman and van DijkはJeffreys型の前分布は「単位根仮説のベイズ分析には不適切である」と批判している。

ベイジアンは古典的な単位根の論文に対して、いくつかの面で批判している(Sims, 1988; Sims and Uhlig 1991; Koop 1992b, c; Koop and Steel, 1994; Schotman and van Dijk, 1991a, b; DeJong and Whiteman, 1991a, b)。単に検出力が低いだけでなく、それらが単位根でのナイフの刃のような漸近性にかかなり依存しているためである。すなわち、古典的な漸近理論は、もし $|\rho| < 1$ なら標準的な結果が成り立つが、 $\rho = 1$ なら結果は大きく異なるという事実に悩まされている。さらに、もし $|\rho| > 1$ で系列が発散するなら、漸近的な結果は単位根あるいは定常系列の場合とは異なる。もちろん、小標本では、定常であるが持続性のある系列(たとえば、 $\rho = 0.95$)から生成された系列は単位根とは実際上区別することができない。これらの論文が示

¹³⁾ しかし、Phillipsの考えたU.S.マクロ経済時系列における非定常性の確率は、2分の1よりもかなり小さいものである。この特性は、何人かの研究者にベイジアン単位根の結果が事前密度にかかなり変化があってもかなり頑健であると結論づけることになる。

しているように、ベイズテストは厳密な有限標本の結果を与える。この方法は ρ について連続な尤度関数に基づいているから、古典的な漸近性における不連続性はない¹²⁾。

ベイズ流単位根検定を行う代表的な2つの方法がある。1つの方法は、 ρ に対する報知情報を用いて、単位根モデルを非単位根モデルと比較する事後オッズを計算する (Koop, 1992b, c; Koop and Steel, 1994; など)。もう1つの方法は報知事前分布を用いず、 ρ が1の近傍の一定の区域 (たとえば、 $P(0.95 < \rho < 1.05 | y)$ あるいは $P(|\rho| \geq 1 | y)$) に入る事後確率を計算する (DeJong and Whiteman, 1991a, b)。これらの考え方と異なる方法が、Tsurumi and Wago (1996) で提案されている。この方法は、DeJong and Whiteman の用いたフラットな事前分布による結果を利用して、事後密度を帰無仮説の値の方向に引っ張る自然共役事前分布を用いて分析している。ここで、次の2つの点を強調できる。1) ベイズ単位根検定は、必要なのは尤度関数と事前分布であるから導出が容易であり、厳密な有限標本の結果を与える。古典的な単位根検定を導出するのに必要な漸近理論は極めて複雑であり、モンテカルロ研究によれば、これらの漸近的な結果を小標本に適用するのは疑問がある。2) ベイズ単位根検定は、古典的な検定よりも単位根の存在をずっと小さく見る傾向にある。

古典的な単位根検定とベイズ単位根検定との差は、用いられるそれぞれ仮説検定の方法の違いに大きく依存している。古典的な検定方法は単位根を帰無仮説の状態で検定し、標本情報が圧倒的に否定する場合だけ単位根を棄却する。逆に、ベイズ事後オッズ比は単位根が存在する確率を計算し、次にそれを、たとえば系列がトレンド定常である確率と比較する。小標本では、単位根モデルから生成されたデータは強い永続性のあるトレンド定常なモデルから生成されたデータと事実上区別できないであろう。強い永続性を示し、時間がたつにつれて上方トレンドがある、あるいはその両方がある有限の標本の系列の場合には、ベイズ法は単位根とトレンド定常モデルが似た事後確率をもつことを明らかにする。しかし、(圧倒的な標本情報に直面する場合だけ棄却する) 古典的な検定方法では、単位根という帰無仮説を棄却できないだろう。別の方法で示すと、ほとんどのマクロ経済時系列が単位根を含んでいるという事実は、古典的な検定が棄却できないことを確実に保証する。ベイジアンはトレンド定常というようなもう1つの仮説よりも単位根 (あるいは非正常性) のほうがどれほど起こりうるかと尋ね、否定的に答えがちである。簡単に言うと、ベイズ法は古典的な受容/棄却という検定方法よりも、標本情報をより適切に要約している。

EBA (Extreme Bounds Analysis: 極値境界値分析)

多くのベイズ分析を用いている人たちはモデル選択の古典的方法を批判しており、たぶん最も強いものは Ed Leamer の研究 (1978, 1983, 1985) から生じたものであろう。Leamer が行っている批判は、経済学は実験科学ではないのだから、**真のデータ生成プロセス**を見つけることができるかと考えることは間違いである、というものである。結果として、主要な推定がモデルの定式化に敏感なのか、あるいは他の仮定に対して敏感なのかを確かめることは重要になる。線形回帰モデルでは、感度分析のパフォーマンスに対するこの一般的な名称は、EBA として知られている一連のテクニックによって適応される。

EBA は明確な実証分析での成功例であるが、これは一部は疑いもなくそれがベイズ分析としても古典的分析としても解釈できるという事実のためである。このことを理解するために、正規誤差をもつ次のような線形回帰モデルを考える。

¹²⁾ 多分、この問題に対するベイジアンと古典派の間の大きな差を最も分かりやすく示したのは、Sims and Uhlig (1991) である。ここでは、ベイジアンと非ベイジアンそれぞれに関心がある $\rho|\bar{\rho}$ と $\bar{\rho}|\rho$ のグラフを提示し、いかにベイジアンのグラフが両者のうちでより合理的かを示している。

$$y_t = x_t' \beta + \varepsilon_t \quad (t=1, \dots, T)$$

ここで x_t は $k \times 1$ の説明変数ベクトル、そして ε_t は $iid. \mathcal{N}(0, \sigma^2)$ である。事前分布は自然共役で、 R と r はそれぞれ既知の $J \times k$ と $k \times 1$ 行列で $\text{rank}(R) = J$ とする。アприオリに $E(R\beta) = r$ を仮定するが、事前分散 V_0 は定式化しないままである。もし、 β の線形結合 ($\phi\beta$, ここで ϕ は $k \times 1$) に関心があれば、この線形結合の論理的推定値は $E(\phi\beta|y)$ となる。しかし、この例では事前分散を定式化できないし、しようとしな。Leamer は $E(\phi\beta|y)$ がとりうる極値を計算することを提唱している。すなわち $\min_{v_0} E(\phi\beta|y)$ と $\max_{v_0} E(\phi\beta|y)$ を解くことを考える。2つの解は、 $E(\phi\beta|y)$ の事後平均がとる極値になる。この差 δ を計算し、パラメータの推定値の曖昧さの尺度としようとするものである。 δ が大きければ、推定値のとりうる範囲が広がることを意味している。言い換えると、EBA は研究者はパラメータの事前の位置についてある情報をもっているが、あえて事前分散行列を定式化しようとしな、共通の場合に対する事前分布の感度分析として解釈できる。

EBA の古典的な解釈も直感に訴え、納得できる。説明変数は2つのグループ、自由変数と候補変数、に分割される。自由変数は回帰に必ず入る変数であり、これらの変数の係数にどんな制約もおかないことに確信を持っている。それに対して、候補変数は推定に必須ではないが方程式の適合度に寄与する変数である。どんな候補変数も $R\beta = r$ の形の任意の線形制約に従うとさせることができる。たとえば、貨幣需要関数では利子率を自由変数として選び、その他のたとえば、金融イノベーションに対する代理変数、ラグ付き変数を、ゼロ制約に従う候補変数とする。Leamer は、これは基本的には多くの応用研究者が通常行っていることであり、通常2, 3のキー変数に関心があり、その他の実験をおこなうと述べている。しかし、この実験から唯一つの最終的な回帰式を示すだけよりも、Leamer は、**重要な変数と考える** キーパラメータのどれかが起こる極値を示すべきであると述べている。簡単には、Leamer は候補変数に課される線形制約に対する重要パラメータ推定値の感度を調べ、これらの極値境界が上のベイズ事前感度問題におけるものと正確に同じになるかどうかを調べる。

EBA はいくつかの実証分析で行われている。たぶん最も重要なものは Cooley and LeRoy (1981) であろう。この研究は McAleer, Pagan and Volker (1985) によって反論され、ふたたび Cooley and LeRoy (1986) と Leamer (1985) によって弁護されている。彼らの研究では、Cooley and LeRoy (1981) は EBA を貨幣需要関数に適用し、利子率弾力性に対する EBA は非常に広く、ゼロを含んでいるので、いかなる推測も不安定なものであるとの結論を得た。たとえば、利子率の符号が負で推定されるのは、候補変数を都合よく組み合わせて得られるものであり、データに含まれている情報だけでは利子率が負である結果は得られない。したがって、推定されている関数の多くは、分析者の強い先験的情報でゆがめられた結果である可能性が大きいことを明らかにしている。McAleer, Pagan and Volker (1985) はこの見解に様々な理由を挙げて反論した。その多くはここで関心はない。今の目的のためには、彼らが EBA が候補変数の有意性に依存しているだけでなく、異なる分類が異なる EBA を意味するという意味で、変数を自由と候補のカテゴリーに分けることに依存している範囲の大きさを強調していることに注意するだけで十分である。これらは Leamer (1985) や Cooley and LeRoy (1986) が厄介な問題とは見なさなかつた性質である。一般的に言うと、これらの性質は十分考慮して遂行された古典的分析とベイズ的な方法を用いた分析との間の本質的な類似性を強調するものと考えることができる¹³⁾。

¹³⁾ EBA の分析例については伴 (1991), 163-165 を参照のこと。

Leamer は SEARCH と呼ばれるコンピュータ・パッケージを作成し、彼が提唱した多くの手法が取り入れられている。

例：単位根のベイズ分析

ARMA 誤差の次数が未知の場合のトレンドモデルにおいて単位根がある場合を示し、系列がトレンド定常プロセス (TSP) と階差定常プロセス (DSP) のどちらであるかを中心に検定する例について、Tsurumi and Wago (1996) に沿って多少詳しく紹介しよう。モデルは、次のように与えられる。

$$y_t = \mu + at + \epsilon_t, \quad \epsilon_t = \rho\epsilon_{t-1} + \frac{\Theta(B)}{\Phi(B)}u_t$$

ここで $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_p B^p$ と $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ である。 $\Theta(B)$ は反転可能で $\Phi(B)$ は定常であると仮定されている。

自己回帰誤差のある回帰モデルで単位根を検定するのに

$$a(B) = \eta B + (1-B)(1 - \phi_1 B - \phi_2 B^2 - \dots)$$

と表し、ここで $\eta = a(1)$ である。このように表し、ラグを p 期で打ち切ると次式のようなになる。

$$\Delta y_t = \mu^* + \eta y_{t-1} + \alpha^* t + \sum_{i=1}^p \psi_i \Delta y_{t-i} + u_t$$

ここで μ^* は新たな定数項、そして $\alpha^* = a\eta$ である。

この式を用いると、DST と TSP 仮説は次のようになる。

$$H_0: \eta = 0 \text{ and } \alpha^* = 0 \text{ vs. } H_1: \eta \neq 0$$

$\delta = (\eta, \alpha^*)$ に対する事後密度を導くために、事前密度を与える必要がある。多くのベイズ分析では非情報フラット事前分布や *Jeffreys* 型の不変事前分布が良好な結果を得る。しかし、単位根のベイズ検定では、フラットな事前分布を用いると $\eta = 0$ の帰無仮説が真であるときに、 η の負の領域に大部分の密度がある事後分布になる。そこで DSP が真であるときに、DSP よりも TSP を受け入れやすい。この点を考慮して、Phillips (1991a) は *Jeffreys* 型の前分布を用いて事後分布を 1 の方に引き寄せる望ましい結果を得ているが、その場合事後分布が双峰分布になり、 $\rho > 1$ の発散区域内では *Jeffreys* 型事前分布は指数的に増大するので裾が厚くなる結果を生む傾向がある。Jeffreys (1961, pp. 358-359) は系列相関の分析に用いる場合、係数が 1 のときにこの *Jeffreys* 型の不変原理を用いる点についていくつかの問題を指摘している。フラットな事前分布を使うときは、 η の最尤推定量 (MLE) $\hat{\eta}$ は、 $\eta = 0$ のとき負になる傾向がある。そこで、単位根あるいは DSP を検定するのに、事後密度を $\eta = 0$ の方向に引っ張る事前密度を選択する必要がある。

そこで、次のような自然共役事前分布を用いた。

$$p(\mu^*, \eta, \alpha^*, \psi_1, \dots, \psi_k, \sigma) \propto \sigma^{-1} p(\eta, \alpha^* | \sigma)$$

ここで

$$p(\eta, \alpha^* | \sigma) \propto \exp\left\{-\frac{1}{2\sigma^2}(\delta - \delta_0)' V_0(\delta - \delta_0)\right\}$$

そして $\delta = (\eta, \alpha^*)$ 、そしてこの事前分布を $\eta = 0$ と $\alpha^* = 0$ の帰無仮説の値の中心におく。 δ の事

後密度を導くために、推定式を次のような行列の形に書きなおす。

$$w = Z\delta + X\beta + u$$

ここで

$$w = \begin{bmatrix} \Delta y_1 \\ \vdots \\ \Delta y_n \end{bmatrix} \quad Z = \begin{bmatrix} y_{-1} & 1 \\ \vdots & \vdots \\ y_{n-1} & n \end{bmatrix} \quad X = \begin{bmatrix} 1 & \Delta y_0 & \cdots & \Delta y_{1-k} \\ \vdots & \vdots & & \vdots \\ 1 & \Delta y_{n-1} & \cdots & \Delta y_{n-k} \end{bmatrix}$$

そして $\beta' = (\mu^*, \phi_1, \dots, \phi_k)$ 。 δ の事後分布は次式によって与えられ

$$p(\delta | \text{data}) \propto [\nu s^2 + d + (\delta - \bar{\delta})(V_0 + V)(\delta - \bar{\delta})]^{-(n-k-1)/2}$$

ここで $V = ZM_xZ$, $M_x = I - X(X'X)^{-1}X'$, $\nu s^2 = w' M_x w$, $d = \bar{\delta}' V_0 (V_0 + V)^{-1} V \bar{\delta}$, $\bar{\delta} = V^{-1} Z' w$, $\bar{\delta} = (V_0 + V)^{-1} V \bar{\delta}$ である。

ここで、2つのことを決めねばならない。(i) ラグの長さ p と、(ii) 事前分布における V_0 の値である。事前密度の V_0 の値を得るために、ブートストラップ法を用いる。

- (1) モデルのパラメータのMLE $\hat{\theta} = (\hat{\mu}^*, \hat{\eta}, \hat{\phi}_1, \dots, \hat{\phi}_p)$ と $\hat{\sigma}$ を用いる。 $\hat{u}_1, \dots, \hat{u}_n$ を $\mathcal{N}(0, \hat{\sigma}^2)$ から抽出し、ダイナミック・シミュレーションを行う。

$$\Delta y_t = \hat{\mu}^* + \hat{\eta} y_{t-1} + \hat{\alpha}^* t + \sum_{i=1}^p \hat{\phi}_i \Delta y_{t-i} + \hat{u}_t$$

計算された Δy_t が利用できるようになると、 y_{t-i} の実際のデータを置き換える。 y_t の $n+r$ 個の値を発生した後、 y_t の最初の r 個の計算値を棄て、残りの n 個の値を用いる。

- (2) y_t の n 個のシミュレートされた値を用いて、新しい $\theta = (\mu, \eta, \alpha^*, \phi_1, \dots, \phi_p)$ を推定する。それらを $\hat{\theta}^{(i)} = (\hat{\mu}^{(i)}, \hat{\eta}^{(i)}, \hat{\alpha}^{*(i)}, \hat{\phi}_1^{(i)}, \dots, \hat{\phi}_p^{(i)})$ とする。

- (3) (1)と(2)を R 回繰り返す。 $\hat{\delta}^{(i)} = (\hat{\eta}^{(i)}, \hat{\alpha}^{*(i)})$ を $i=1, \dots, R$ に対して集め

$$C = \text{Cov}(\delta) = \frac{1}{R} \sum_{i=1}^R (\hat{\delta}^{(i)} - \bar{\delta})(\hat{\delta}^{(i)} - \bar{\delta})'$$

を計算する。ここで $\bar{\delta} = 1/R \sum_{i=1}^R \hat{\delta}^{(i)}$ 。さらに $V_0 = C^{-1}$ とする。

R 回の抽出のそれぞれに対してMLEである $\hat{\theta}$ と $\hat{\sigma}$ を用いる代わりに、 $i+1$ 回目の抽出に $\hat{\theta}^{(i)}$ と $\hat{\sigma}^{(i)}$ を用いることもでき、これはまさにGibbsサンプラーの方法である。 t と y_{t-1} のような強いトレンドがある説明変数の場合には、 y_t の拡張されたデータが大きく変動するためにGibbsサンプラーが難しい場合がある。

いくつかのベジアン・ブートストラップ法が提案されているが、ブートストラップ法は用いるモデルとデータを反映するように計画されており、古典派とベイズ派の折衷的な弾力性を持っている。 η と α^* の事後分布も、簡単に計算することができる。

Tsurumi and Wago では、為替レートモデルでこの単位根検定を行っている。DSP対TSP仮説を検定するために、最高事後密度区域 (HPDR) あるいは最高事後密度区間 (HPDI) 検定を行う。もしHPDR検定なら α_1 の所与の水準 (たとえば $\alpha_1 = 0.05$) に対して、 $(1-\alpha_1)$ HPDR 等高線が点 $(\eta=0, \alpha^*=0)$ を含むか否かを見る。HPDI検定の場合には、個々に η の $(1-\alpha_1)$ HPDI が $\eta=0$ を含み、 α^* の $(1-\alpha_1)$ HPDI が $\alpha^*=0$ を含むかどうか見る。単に点推定値や検

定統計量を見るだけよりも同時事後分布や周辺事後分布をプロットすると、 η と α^* についてもっと多くの情報を得ることができる。

次の為替レート決定モデルを考える。

$$s_t = \beta_1 + \beta_2(m_t - m_t^*) + \beta_3(y_t - y_t^*) + \beta_4(r_t - r_t^*) + \beta_5(p_t - p_t^*) + \beta_6 f_t + u_t$$

ここで、次の各変数は時間 t での対数をとったものである： s_t =名目為替レート、 m_t =貨幣供給、 y_t =生産、 r_t =短期利子率、 p_t =卸売物価指数、 f_t =外国為替準備、である。なおアスタリスク(*)は外国での数量を表している。

この式に含まれている6変数について、各変数が同じ次数で和分されているかどうか検定した結果が次の表である。この表には、方程式の η と α^* に対してDSP対TSP仮説を検定するために、Dickey-FullerのF検定が95%最高事後密度区間(HPDI)とともに報告されている。

Dickey-Fuller F-検定と最高事後密度区間(HPDI)

変数	ADF F-統計量	η に対する 99% HPDI	α^* に対する 99% HPDI
s_t	3.98	(-.07, .025)	(-.0004, .0002)
$m_t - m_t^*$	3.84	(-.085, .025)	(-.0003, .0001)
$y_t - y_t^*$	3.94	(-.135, .030)	(-.0001, .0003)
$r_t - r_t^*$	2.64	(-.050, .021)	(-.0002, .0003)
$p_t - p_t^*$	7.15	(-.075, .009)	(-.0002, .0001)
f_t	4.15	(-.050, .015)	(-.0001, .0004)

ここで、99% HPDIは99%最高事後密度区間を表している。またラグの長さは12($p=12$)としている。

表から、Dickey-Fuller F検定は5%有意水準で $p_t - p_t^*$ 以外のすべての変数はDSPであることが分かる。 $p_t - p_t^*$ に対しての7.15というF統計量は5%水準で有意であるが、5%水準からそんなに離れていない。他方、99%-HPDIは $\eta=0$ と $\alpha^*=0$ の帰無仮説の値の回りに集中している。従って、6つの変数がすべて単位根があることになる。

4. ベイズ法における計算

第2章で強調したように、事後平均とか事後オッズのような統計量を解析的に求める方法は滅多に存在しない。そのような場合には、ベイズ流の応用研究では数値的な手法に頼らなければならない。そこで、そのいくつかの方法について述べる。まず最も単純な近似から始め、次にもっと手の込んだ方法を考えることにする。

事後分布を求めるには、複雑で、しばしば高次元の積分を評価することが必要になる。加えて、事後分布の要約統計量の計算にはモーメントや四分位値が含まれており、そのためにはさらに積分計算が必要になる。この問題に対する初期の取り組みは事後密度への解析的な近似解を得るために漸近的な方法を用いることであり、最も単純な結果は事後分布への正規近似で、基本的には中心極限定理のベイズ版といえる。もっと複雑な漸近的な方法はラプラス法(Tierney and Kadane, 1986)の様な手法で、もう少し正確で、かなり非対称な事後分布への近似ができる。このような近似法が適用できないかあるいは精度が不十分である場合には、数値積分に頼ることになる。何年もの間ガウス求積法のような伝統的な方法がこの目的のために使われてきた¹⁴⁾。しかし、この方法を適用するのは、対象とする計算式が大きなモデルでは扱えないので、(パラメータ数が10位までの)低次元の計算に限られてきた。これらの方法には、一定の

精度を維持するために必要な関数を評価する数が、次元数とともに指数的に増加するという次元の呪いという問題がある。もう一つの方法は、Dempster, Laird and Rubin (1977) の EM アルゴリズムである。本来は欠測値問題における最尤推定値を見つける道具として開発された方法であるが、多くのベイズ分析での事後分布のモードを見つけるのに用いられている。EM アルゴリズムは多くの実際問題に適用されているが、かなり遅いアルゴリズムであり、事後分布全体を推定するよりも主として事後モードを見つける目的だけに向けられている。結果として、最近の多くの応用分析家はモンテカルロ積分法に注目するようになった。この方法はより完全な情報を提供し、かなり高次元のモデルでも比較的程序が簡単であるという特徴がある。この方法には、インポータンス・サンプリング (Geweke, 1989a) や単純な受容/棄却サンプリングのような伝統的な非反復的方法が含まれる。もっと強力な方法は、最近開発された反復的マルコフ連鎖モンテカルロ法で、データ拡大 (Tanner and Wong, 1987)、メトロポリス-ヘイスティングス・アルゴリズム (Hastings, 1970)、そしてギプス・サンプラー (Geman and Geman, 1984; Gelfand and Smith, 1990) が含まれる。これらの方法はマルコフ連鎖を作成し、そのアウトプットは同時事後分布からの (相関のある) サンプルに対応する。これらの方法について、以下順に紹介する。

4.1 正規近似

多くの場合、事後密度は近似的に正規分布にしたがう。正規分布の性質はよく知られているので、事後密度から近似的な推測を行うのは容易である。説明上、そのモードの回りで事後密度の対数をテーラー展開する。

$$\ln p(\theta|y) = p(\hat{\theta}|y) + (\theta - \hat{\theta})' G(\hat{\theta}|y) + \frac{1}{2}(\theta - \hat{\theta})' H(\hat{\theta}|y)(\theta - \hat{\theta}) + \text{残余}$$

ここで $G(\theta|y)$ は θ で評価した傾斜、 $H(\theta|y)$ は θ で評価したヘッセ行列である。モードで評価した傾斜はゼロであることに注意し、残余を無視すると、事後密度は近似的に正規分布にしたがうことが分かる。実用的には、これは近似的なベイズ分析はちょうど古典的な計量経済分析で最尤値を求めることに対応している。標準的な最適化のパッケージであれば事後モードを得るのに使うことができ、対数事後密度のヘッセ行列の逆行列にマイナスを付けたものを近似的な事後共分散行列として用いる。たとえば、Zellner and Rossi (1984) はロジットモデルでの研究にこの近似を採用している。しかし、彼らは正規近似が極めて良くないことを見出している。

正規近似に関する問題は、近似誤差が非常に大きいことである。それはこの方法が事後平均に実際は事後モードで近似しているためである。もし、事後分布がかなり歪んでいるか双峰ならば、結果は惨めである。また、パラメータの非線形関数である量に関心がある場合は、正規近似はそれほど役に立たない。さらに、この方法は近似の精度を評価するための診断が得られない。可能なら、次に述べる方法を用いてもっと正確な分析をするべきであろう。

4.2 数値積分

ベイズ実証分析に通常見られるタイプの積分を計算する最も直接的な方法は、多くの数値積分公式の中の1つを用いることである。これらの手法の中で最も単純なものはシンプソンの公式を用いており、多少手の込んだものはガウス・エルミート球積法を用いる。これらの手法は、標準的な汎用数値積分プログラムが低次元の問題に対してだけ有効であるという重大な欠点がある。しかし、計量経済分析では通常多次元の問題を取り扱う場合が多い。Naylor and Smith

¹⁰ これらの方法についての一般的な議論については Thisted (1988, Ch. 5) を、また事後分布の計算への応用については Naylor and Smith (1982) を参照。

(1982, 1988) と Smith et al. (1985) は数値積分の全般的な議論をしており、9次元までの数値積分を行っているが、プログラムは複雑で、個別の問題に対するものである。経験のあるコンピュータ・プログラマーだけが使いこなせるものである。しかし、1つか2つのパラメータに関する積分が必要な応用問題に対しては、数値積分は実行可能である。たとえば、BRAP¹⁵⁾は、1次元の問題に対してはシンプソンの法則に基づいた簡単なサブルーチンがあり、IMSL や NAG ライブラリーは Fortran プログラム¹⁶⁾から直接呼び出せるいくつかのサブルーチンを含んでいる。多分最も手の込んだプログラムは Bayes Four であり、数値積分を行うための広範囲のサブルーチンを含んでいる (Naylor and Shaw (1991) 参照)。

4.3 ラプラス近似

ベイズ分析では積分を近似するためにラプラス法を用いることができる。Tierney and Kadane (1986) と Tierney, Kass and Kadane (1989), そして Kass, Tierney and Kadane (1988) による研究によって、ベイズ計算に役立つ解析的な近似を得ることができるようになった。特に、実数値あるいはベクトル値パラメータや周辺事後密度の事後平均と分散を近似する積分にラプラス法を適用した。この方法を用いると、ベイズ分析は積分 $\int f(\theta)\exp[-Th(\theta)]d\theta$ (スカラーの θ に対して) を

$$f(\theta)\sqrt{\frac{2\pi}{T}}\sigma\exp[-Th(\hat{\theta})]$$

によって近似する。ここで θ は $h(\theta)$ の最大値、そして σ^2 は θ で評価したヘッセ行列の逆行列にマイナスを付けたものである。ほとんどのベイズ積分はこの形を取る。たとえば Tierney and Kadane (1986) は

$$E(g(\theta)|y) = \frac{\int \exp(TL^*)d\theta}{\int \exp(TL)d\theta}$$

と表している。ここで $L = [\ln(p(\theta)) + \ln(L(\theta|y))]/T$, $L^* = L + \ln[g(\theta)]/T$ として T は標本のサイズである。ラプラス法を分子と分母に適用すると、両者とも非常に似ているので近似誤差も非常に似ており、そこで互いに相殺することが分かる。結果的に、Tierney-Kadane 法を用いて計算された積分は通常極めて正確である。

多変量の場合には、次のように $E(g(\theta)|y)$ に対する 2 次近似値を得る。

$$\hat{g} = \frac{\Sigma^*(\theta^*)^{1/2} \exp[TL^*(\hat{\theta}^*)]}{|\Sigma(\theta)|^{1/2} \exp[TL(\hat{\theta})]}$$

ここで θ と θ^* は最大値であり、それぞれ L と L^* に対応する Σ と Σ^* は、ヘッセ行列の逆行列にマイナスを付けたものである。Kass, Tierney, and Kadane (1989) で示されているように、この近似の相対誤差は T^{-2} のオーダーで上の正規近似よりもかなり良い。対応する $g(\theta)$ の分散の 2 次近似は次のようになる。

$$\widehat{\text{Var}}(g(\theta)|y) = \widehat{E}(g(\theta)^2|y) - [E(g(\theta)|y)]^2$$

Tierney-Kadane 法を行うには、関数を最適化することだけが必要である。特に L と L^* は最

¹⁵⁾ BRAP はシカゴ大学の Arnold Zellner を中心としたグループによって考案された TSP に似たベイズ計量経済分析用のコンピュータパッケージである。このプログラムの詳細は付録を参照のこと。

¹⁶⁾ IMSL と NAG ライブラリーは、ほとんどの大学で利用できる Fortran サブルーティン・ライブラリーである。

大化され、そのヘッセ行列を得る必要があるが、多くの最適化パッケージが存在するのでわざわざプログラムする必要はない。最尤法をよく使う応用分析を行う人にとっては、積分問題を2つの最適化問題に変えるので Tierney-Kadane 法は簡単に使える。そして L と L^* は似た関数であるから、もし $\hat{\theta}$ を $\hat{\theta}^*$ の探索での初期値として使えば計算が容易になる。Tierney-Kadane 法は、たとえ $\hat{\theta}^*$ が $\hat{\theta}^*$ の方向で $\hat{\theta}$ からの1ニュートン・ステップによって置き換えられても、その高い精度を漸近的に維持する。この方法は、Tierney-Kadane の1段階法として知られている。簡単には、(いくらか精度を犠牲にしても) 2段階目の最適化を行う必要はない。

Tierney-Kadane 法に関しては、4つの欠点を指摘しておく必要がある。第1に、この方法は計算したい特性値に対して繰り返さなければならない。したがって、もし20個のパラメータの事後平均と分散を計算する場合には、41回の通常時間がかかる最適化を行わなければならない。第2に、この方法は $g(\theta|y)$ の対数をとるので正の関数に限定される。しかし、この制約は大きな定数を $g(\theta|y)$ に加えるか、(常に正である)積率母関数を用いてそれを微分することによって回避することができる。第3に、この方法は、近似の精度を診断する方法がない。第4に、この方法の近似の質を改良する方法がない。事実、たとえ Tierney-Kadane 近似の精度が当面の応用問題にとってはあまり良くないということに気付いていたとしても、全く別の方法を考えることができない。

まとめると、Tierney-Kadane 法は少数の重要な積分を速く近似する非常によい方法である。しかし、10個以上の多数の積分を計算しなければならないような大きな問題に対しては、モンテカルロ積分の方が好ましい方法である。

例：ロジットモデル

Koop and Poirier (1993) は、1982年の社会調査を用いて337人の雇用されている男性に対して、5つの職業カテゴリー間での選択が3つの特性に依存している多対ロジットモデルを考えた。次の多対ロジット尤度が仮定される。

$$L(y|\theta) \propto \prod_{i=1}^T \prod_{j=1}^5 P_{ij}^{y_{ij}}$$

ここで、個人 i が職業 j を選択すると $y_{ij}=1$ (その他の場合は $y_{ij}=0$)、そして $P_{ij} = \text{Prob}(y_{ij}=1)$ である。 P_{ij} は、次のように定式化される。

$$P_{ij} = \frac{\exp(z_i \theta_j)}{\sum_{j=1}^5 \exp(z_i \theta_j)}$$

ここで、 z_i は個人 i についての 4×1 のデータベクトル(経験、教育、人種の3つの特性値と定数項)、そして θ_j は 4×1 のパラメータのベクトルである¹⁷⁾。この尤度関数をフラットな事前分布と結びつける。すなわち、事後密度は尤度関数に比例する。目的はいくつかの特性を持つ個人が職業につく確率を予測することである。言い換えれば、 f を特性 z_f をもつ個人の指標とした場合に、 P_{ff} を計算しようとした。これらの予測された確率は、ロジットパラメータ θ の非線形関数である。

最尤推定量に対して closed form の解は存在しないことはよく知られているので、最尤推定値を計算するためには最適化プログラムが必要になる。さらに、一般にこの尤度関数を含む積分に対する解析的な解は存在しないことは容易に証明できる。

¹⁷⁾ θ_j はゼロに正規化されている。

そこで、予測された確率の事後密度を評価するのに、Tierney-Kadaneの方法とその1段階法(One-Step methods)の両方を用いている。前者の方法では、ロジット事後密度の対数(Tierney-Kadaneの記号では L)を最適化し、次に L^* のその後の最適化に対する初期値としてこの最適化の値を用いている。 L^* は計算された各事後平均、すなわち $E(P_{ij}|y)$ で異なるから、多くの最適値が必要になる。モンテカルロ積分の結果と比較すると、これらの両方の近似の精度はこの例では非常によく、必要な計算は最小ですむ。

5. モンテカルロ積分

数値解析の最も簡単な方法は、モンテカルロ・サンプリングに基づく方法である。モンテカルロ積分は、ベイズ実証分析で多分最も広く用いられている手法である。計量経済学への導入は、ほとんどKloock and van Dijk (1978) と Geweke (1989) の研究を通じて行われている。モンテカルロ・サンプリングは積分を直接に数値的に評価するか、あるいは数式を用いて解析的に求めた解の代わりに妥当な標本空間における多数の点で構成される確率分布に依存している。

この方法は、次のようないくつかの利点がある。それらは1)高次元問題を扱うのに適している、2)目標に近似する際の誤差のサイズをコントロールできる、3)プログラムが簡単である。しかし、モンテカルロ法を用いた積分計算がコンピュータ時間の大半を占める場合が多い。

5.1 単純なモンテカルロ積分

単純なモンテカルロ積分は、もし n 個の確率変数 $(\theta^i, i=1, \dots, n)$ を事後分布から抽出すると

$$g^* = \frac{1}{n} \sum_{i=1}^n g(\theta^i) \rightarrow E(g(\theta)|y) \quad n \rightarrow \infty$$

となるという事実に基づいている。この単純な定理に基づくコンピュータ・プログラムは非常に簡単に書くことができる。必要なことは、事後分布から繰り返し乱数を抽出し、これらの乱数それぞれに対して必要な統計量を計算し、その各統計量を抽出回数で平均することである。近似の精度は n に依存し、それをコントロールすることができる。 n についての漸近理論を用いると、数値的な精度の尺度を導くことができる。すなわち、弱い正則条件の下で、 g^* は漸近的に、平均 $E(g(\theta)|y)$ と分散 σ^{*2}/n の正規分布にしたがう。ここで、 $\sigma^{*2} = E(g(\theta)^2|y) - [E(g(\theta)|y)]^2$ である。 σ^{*2} は

$$s^{*2} = \frac{1}{n} \sum_{i=1}^n g^2(\theta^i) - g^{*2}$$

によって推定することができる。 $E(g(\theta)|y)$ に対する95%信頼区間は、 $[g^* - 1.96s^*/\sqrt{n}, g^* + 1.96s^*/\sqrt{n}]$ となる。ここで s^*/\sqrt{n} を数値標準誤差(NSE)と呼ぶ。漸近理論によれば、 $g(\theta)$ の標準誤差のNSEに対する比は $1/\sqrt{n}$ を意味する。したがって、もし $n=10,000$ なら、漸近理論からNSEはその関数の事後標準偏差の1%になることを意味する。

対称反復法

数値近似の効率を増加するために、モンテカルロ・サンプリングを僅かに変形するのが対称反復法である。もし事後密度が近似的に対称であるなら、一定の精度を達成するのに必要な抽出回数は、Geweke (1988) の提案した対称加速法(*antithetic acceleration*)を使えばかなり減らすことができる。この方法は確率分布のモーメント自身の数値評価に対する方法ではなく、標準的なモンテカルロ・サンプリングに加えられる方法である。この手法は次のようにペアに

してサンプルを抽出する。モードの一方でのそれぞれの抽出と比例して、逆の方向で1つ抽出する。もっと正確には、もし事後分布のモードが μ で i 番目の抽出が θ^i であったとすると、対称加速法は $(i+1)$ 番目の抽出を $2\mu-\theta^i$ とする。Geweke (1988) によれば、この方法は数値的な精度を大きく増加させ、したがって、特に事後分布への漸近的な近似が正確ならば n を大きく減少させることができる。このように対称反復法は、確率分布やそのモーメント自身の数値評価に対する一つの方法ではなく、標準的なモンテカルロ・サンプリング図式に加えることができる方法である。対称反復法を導入するには、前述の基本プロセスに1ステップを加え、事後分布内の点が分布の平均の回りに公平に分布する様にする。正の分布から確率抽出された後で、対称反復法は平均を通じて投影した直近の抽出のミラーイメージを作成する。このより早い率で真の値へ収束する数値近似によって数値近似の分散の減少は、サンプルサイズがかなり大きくなると計算効率がかなり向上する。標準的なモンテカルロ・サンプリングで適切な数値精度を保証するのに $n=5000$ と $n=10,000$ が通常選択されるが、もしデータセットが $T=100$ のサンプルサイズなら、500と1000の間の対称確率抽出(250から500組)で一般に同じ程度の数値精度が達成できる。対称反復での数値精度の正確な増加は、関数 $g(\theta)$ の非線形性の程度に依存し、非線形性が増加するにつれて利点が減少する。

例：単位根検定

時系列モデルでは、回帰パラメータの関数にしばしば注目する場合がある。例を用いて、DeJong and Whiteman (1991a, b) で述べられているベイズ単位根検定の方法を考える。DeJong and Whiteman は次のモデルを使っている。

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \beta_4 t + \varepsilon_t$$

ε_t は平均がゼロで分散が σ^2 の*iid.*正規、 $\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)$ 、そして非報知事前分布 $p(\beta, \sigma^2) \propto \sigma^{-2}$ が仮定されると、 β に対する周辺事後密度は t 分布に従う。したがって、 β についての推測は多変量 t 密度の性質を用いて極めて簡単に行える。しかしながら、DeJong and Whiteman は β 自体には関心がなく、自己回帰のdominate(支配的)な根 Λ に関心がある(Λ は自己回帰表現で絶対値での最大の根として定義される)。DeJong and Whiteman は $p(\Lambda \geq 0.975|y)$ のような数値を計算した。残念ながら、これらの数値は β の高度な非線形関数であり、したがって解析的な結果は存在しない。しかし、 $p(\beta|y)$ からのサンプリングは簡単であるから、簡単なモンテカルロ積分を用いることができる。

事実、 t 事後分布は、1) β の最小2乗推定値である平均、2)最小2乗推定から得られる統計量である分散、そして3)自由度、すなわち $T-1$ で決められる。これらの統計量はどんな計量経済分析用のコンピュータ・パッケージからでも簡単に得ることができ、乱数は t 事後密度からいくつかの方法で得ることができる。BRAPのようなプログラムでは、もし上の1)、2)、3)を与えれば自動的に t 密度から乱数を得ることができる。別の方法として、応用経済分析を行う人は、(たとえば、GAUSSのような)ほとんどのパッケージに組み込まれている正規乱数発生機能を使って t 乱数を簡単に作成することができ、次に正規分布と t 分布との関係を利用して t 密度に変換する。

t 乱数の場合では、モンテカルロ積分は事後分布から乱数 β^i を発生し、 Λ を計算する(1回毎の抽出から Λ を計算する)。 Λ に関しての平均は $E(\Lambda|y)$ に収束し、0.975以上の Λ^i の割合は $p(\Lambda \geq 0.975|y)$ に収束する。

DeJong and Whiteman は、単位根の存在を証明するのに $p(\Lambda \geq 0.975|y)$ のような値を使っている。しかし、モンテカルロ積分は自己回帰モデルやベクトル自己回帰(VAR)モデルでのその他のいろいろな統計量を計算するのに使うことができる。もし事前分布が自然共役あるいはフ

ラットであれば、事後密度は多変量 t あるいは行列変量 t になり、事後密度から乱数を抽出することができる。VAR に対するこの種のテクニックは、Doan (1988) と Koop (1992) を参照すると良い。

5.2 インポートランス・サンプリングを用いるモンテカルロ積分

乱数が事後分布から簡単に得られない場合は、インポートランス・サンプリングを用いたモンテカルロ積分が必要になる。単純なモンテカルロ積分は事後分布から直接乱数を抽出し、次に抽出した数で単純平均をとる方法であるが、他方、インポートランス・サンプリングを用いるモンテカルロ積分は、インポートランス関数と呼ばれる代替密度から乱数を発生し、次にその抽出された乱数の加重平均をとる方法である。

たとえば、事後密度の事後期待値を近似する場合を考える。

$$E(f(\theta)|y) = \frac{\int f(\theta)L(\theta)p(\theta)d\theta}{\int L(\theta)p(\theta)d\theta}$$

正規化された尤度に事前密度を掛けた $cL(\theta)p(\theta)$ を簡単にサンプルできる、たとえば、多変量 t 密度やスプリット t のような密度¹⁸⁾ $g(\theta)$ によってラフな近似をすることができる (Geweke, 1989a を参照)。ここで $g(\theta)$ は $cL(\theta)\pi(\theta)$ に近ければそれだけ近似がよい。他方、 $g(\theta)$ への近似が悪ければ多くのウェイトはゼロに近くなり、僅かな θ^i によって合計が変動し、不正確な近似になる。

$g(\theta)$ をインポートランス関数とし、それから簡単に乱数を発生することができる。 θ^i を $g(\theta)$ からの乱数とし、ウェイト関数を $w_i = p(\theta^i)L(\theta^i|y)/g(\theta^i)$ とする。非常に弱い条件の下で¹⁹⁾

$$g^* = \frac{\sum_{i=1}^n w_i f(\theta^i)}{\sum_{i=1}^n w_i} \rightarrow E(f(\theta)|y) \quad n \rightarrow \infty$$

を示すことができる。この定理は一見するとすべてのインポートランス関数に対して成り立つように見える。しかしながら、現実にはもし適切なインポートランス関数が選択されなければ、 n はかなり大きくなる。この選択に対する良い経験則は、インポートランス関数を事後密度関数に近似するか、少なくとも事後分布より裾を厚くとるべきであるということである。もし事後分布の裾がインポートランス関数のそれよりも厚ければ、推定は非常に不安定である。それは、インポートランス関数の裾のずっと端の方でサンプリングされたときを考えると分かる。もし事後分布の裾がインポートランス関数のそれよりも厚かったら、このサンプリングに対する w_i は非常に大きくなり、したがって推定値に非常に大きなウェイトを与える。言い換えると、裾の方での滅多にないサンプリングが $E(f(\theta)|y)$ のモンテカルロ推定値の動きを支配する傾向になる。

ベイズ分析で共通によく使われるのは、前述の t インポートランス関数である。 t 分布は3つのパラメータをもっており、平均、分散、そして自由度である。関数の最適化は比較的易しいので、インポートランス関数のパラメータを選択するためには、事後分布の対数を最適化すればよい。ほとんどの最適化パッケージは自動的に事後モードを計算し、モードでのヘッセ行列を計算する。インポートランス関数の平均と分散は、それぞれ事後モードと対数事後分布のマイナスのヘッセ行列の逆行列に等しいと置く。ほとんどの応用分析で、これはインポートランス関数

¹⁸⁾ スプリット- t 密度は、非対称な場合に認められる以外は通常の t と同じである。定式上は、モードの一方での分散は反対側では異なる場合がある。

¹⁹⁾ 基本的には、これらの条件は $g(\theta)$ のサポートが $p(\theta|y)$ のサポートをカバーしていることが必要である。

の形状が事後分布のそれと似たものであることを保証する。自由度のパラメータは、インポータンス関数の裾が事後分布のそれよりも厚いことを保証するために、裾の動きをコントロールするのに使うことができる。

ところで、 t 分布が事後分布の十分な近似にならない場合には、解析的な方法かあるいは等高線プロットのような数値的な道具のどちらかを用いて事後分布を注意深く調べるしかない²⁰⁾。しかし、多くの応用研究では事後分布は近似的に正規であり、正規あるいは t インポータンス関数を用いると十分満足するものが得られる。さらに、対称加速法を使うことによって必要なサンプルの数を減少させることもできる。

ひとたび1つのインポータンス関数が得られれば、その質を評価してモンテカルロ計算の実行毎の推定値の安定性を調べるための最も単純なアドホックな方法が必要になる。すなわち、10,000個のサンプリングを行う2つのモンテカルロ計算を行って、それらが非常に似た推定値を得たなら、これがサンプリングの数が高次の数値的な精度を達成するのに十分である良い証拠になる。同様に、事後分布に十分よく近似するインポータンス関数は、大体コンスタントになるウェイト w_i になる。そこで、ベイズ分析でインポータンス関数が適切であるかどうかは、1) ウェイトを調べる、2) これらの変動係数を計算する、3) これらのウェイトの分布のグラフをプロットする (Hop and van Dijk, 1990)、のいずれかを行う。

あるいはもう1つの別の方法として、Geweke (1989) による漸近理論を用いたインポータンス関数の診断を行うことができる。前節で紹介した Geweke が用いた最も単純な診断法は、数値標準誤差 (NSE) である。この場合 NSE の推定値は、次式で与えられる。

$$NSE = \sqrt{\frac{\sum_{i=1}^n [g(\theta^i) - g^*]^2 w_i^2}{\left[\sum_{i=1}^n w_i\right]^2}}$$

もう1つの有望な数値は相対的数値有効性 (RNE) であり、Geweke は次式で推定した。

$$RNE = \frac{\text{Var}(g(\theta)|y)}{n^*NSE^2}$$

$\text{Var}(g(\theta)|y)$ はモンテカルロ法によって通常は計算されるので、RNE は NSE を一度計算してしまえば簡単に得ることができる。RNE の大きい値が望ましい。すなわち、もし $RNE=1$ なら、インポータンス関数は事後分布から直接サンプリングして得た結果と同程度の精度を持つ結果を得る。さらに、NSE は $(RNE \cdot n)^{-1/2}$ と事後標準偏差の積である。したがって、 $RNE=0.1$ で $n=1,000$ とすると、数値標準誤差は事後標準誤差の10%になる。

事後オッズの計算には、 $E(f(\theta)|y)$ の形にはできない $\int \pi(\theta)L(\theta|y)d\theta$ の計算が必要である。しかし、Geweke (1989) は次式になることを示した。

$$\int \pi(\theta)L(\theta|y)d\theta \doteq \frac{1}{n} \sum_{i=1}^n w_i$$

すべての正規化定数は事後オッズを計算する場合には含める必要がある。あるいは、少なくとも分析の最後で再び加える必要がある。すなわち、 $\int p(\theta)d\theta=1$, $\int \ell(\theta|y)d\theta=1$, そして $\int g(\theta)d\theta=1$ が必ず保証される。

²⁰⁾ たとえば van den Broeck, Koop, Osiewalski and Steel (1992) の研究では、事後分布の詳しい分析によってガンマ・インポータンス関数が適切であることが示されている。

インポートランス・サンプリングの計算

インポートランス・サンプリングは、ガウス型尤度関数を持つが共役でない報知事前分布をもつ共通の回帰モデルを分析するのに有効な方法である。そこで、事後密度は非正則で、そこから確率抽出が困難か不可能である。インポートランス・サンプリングの計算アルゴリズムは、まとめると次の通りである。 $f(\theta|y)$ を y が与えられたときの条件付き確率密度関数としたとき、もし $p(\theta)$ から確率変数を簡単に発生することができるなら、 $h(\theta) = \int f(\theta|y)p(\theta)d\theta$ を求めるために

1. $p(\theta)$ から互いに独立な n 個の標本 y_1, \dots, y_m を発生させ
2. $\widehat{h(\theta)} = 1/m \sum_{i=1}^m f(\theta|y_i)$ を計算する

$h(\theta)$ からの標本を発生させるには、 $\widehat{h(\theta)}$ の代わりに $f(\theta|y_i)$ から θ_i を発生させる。しかし、 $p(\theta)$ から標本を発生させることが難しい場合には、この方法は適用できないのでGeweke(1989)は $p(\theta)$ に近い、標本を発生させやすいインポートランス分布と呼ばれる分布 $g(\theta)$ を用いて

- (1) $g(\theta)$ から互いに独立な m 個の標本 y_1^*, \dots, y_m^* を発生させ
- (2) $\widehat{h(\theta^*)} = \sum_{i=1}^m w_i f(\theta|y_i^*) / \sum_{i=1}^m w_i$, $w_i = p(\theta_i^*)/g(\theta_i^*)$ を計算する

とすると、 $g(\theta)$ のサポートが $p(\theta)$ のサポートを含むとき、確率1で $\widehat{h(\theta^*)}$ が $\widehat{h(\theta)}$ に収束することを証明した。 $h(\theta)$ からの標本を発生させるには(2)の代わりに

- (2)' $f(\theta|y_i^*)$ から θ_i を発生させる ($i=1, \dots, m$)
- (3) $(\theta_1, \dots, \theta_m)$ に確率 (w_1, \dots, w_m) を与えて、標本 $(\theta_1^*, \dots, \theta_m^*)$ を発生させる

とすればよい。Smith and Gelfand (1992)はこれを重み付ブートストラップと呼び、 m が大きくなるに従って $h(\theta)$ から発生される確率標本に対する近似がよくなることを証明した。

インポートランス密度の選択は、数値近似と計算効率の両方で重要である。2つの密度が似ていれば、それだけ数値近似がよくなる。インポートランス密度の選択で最も重要な点は、インポートランス密度の裾のカバレッジが少なくとも真の事後分布の裾のカバレッジに等しいか、そうでなければ少しだけ大きいことを保証することである。すなわち、インポートランス密度は θ の事後密度よりはいくらか裾が厚い。これによって、事後密度の裾の方では標本をとりすぎることになるが、これらの観測値はインポートランス・ウェイト関数 $w(\theta)$ によって小さくウェイト付けられる。もし裾が非常に薄ければ裾のカバレッジが不適切であり、これらの観測値に過剰にウェイトを与え、数値サンプリングをおこなう時に標本空間の全領域を完全に覆えない(Geweke, 1989a)。

そのために、通常適切なカバレッジを保証するためにスケールアップした分散・共分散行列のインポートランス密度を用いる(Tanner, 1996, p. 56)。たとえば、 n 個の観測値と k 個の外生変数のある単一方程式回帰モデル $y = X\beta + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$ を仮定し、ここで $\theta = (\beta, \sigma^2)$ に対する事前分布は $h(\beta)/\sigma$ に比例する。すると、回帰係数 β には報知事前情報を、そしてスケールパラメータには標準的な散漫な事前分布をもつ。データの尤度は標準多変量正規分布であり、 $h(\beta)$ が一定なら事後分布も $\theta = (\beta, \sigma^2)$ に対して多変量正規-逆ガンマ型であるような標準的な形になる(Zellner, 1971)。しかし、多くの経済分析への応用の場合では、 $h(\beta)$ は β の標本空間を越えて一定であることは滅多にない。経済理論によってある点で切断されるかもしれないし、正の値だけしかとらえないかも知れない。また、以前に行われた研究から得られた知識によって特別な情報を持つこともある。多くの場合、インポートランス・サンプリングは β あるいは $g(\beta)$ の事後分布を調べるために用いる最適な方法である。データの尤度関数は、事前分布の形によ

って尤度関数や事後分布が（事前分布のサポートがある標本空間の区域内で）極端に異なるものではないと仮定すると、しばしばインポートانس密度は自然に選択される。 (β, σ) についての確率抽出は簡単に行うことができ、したがって事前分布は評価することができ、前述の全ステップで完全な分析ができる。

計算ソフトウェア問題

インポートانس・サンプリングを用いたモンテカルロ積分を行ういくつかのパッケージプログラムがある。たとえば、BRAP は正規あるいは t インポートانس関数を用いてモンテカルロ積分を行うプログラムである。多分現在利用可能な最も手の込んだモンテカルロ積分パッケージは、Hop and van Dijk (1990) によって開発されたものであり、2つのアルゴリズムでプログラムした SISAM と MIXIN である。SISAM は、 t インポートانس関数を用いたモンテカルロ積分を行う。他方、MIXIN はもっと複雑な混合積分を用いて、モンテカルロと数値積分の両方を結合して歪んでいる可能性のある密度を含む積分の正確な評価を行う。両方のサブルーティンを使うと、広範囲のインポートانس関数の診断ができる。BRAP のモンテカルロ積分の機能を利用するには Fortran が必要である。Fortran と NAG サブルーティン・ライブラリーが SISAM と MIXIN を利用するのに必要である。

しかし通常は GAUSS, S-Plus や Fortran, あるいはどれか共通のプログラム言語を用いて特定目的のプログラムを開発する場合が多い。このとき必要なのは、サンプリング毎に事後分布を評価するサブルーティンである。すなわち、インポートانس関数から標本を抽出し、そのサンプリング²¹⁾でのインポートانس関数を評価するサブルーティンと、すべての機能を含んだメインプログラムの部分は各サンプリングで評価し、適切にウェイト付けする。インポートانس関数に対するパラメータ値を得るために、事後分布を最初に最適化する必要がある場合もある。この場合では、GAUSS のようなパッケージは最適化サブルーティンをもっており、ユーザーはすでに最適化する関数に関してはプログラムしている。

受容/棄却法

インポートانس・サンプリングで $p(\theta)$ から標本を発生させにくい場合には、次のような受容/棄却法 (A-R: *Acceptance-Rejection Sampling*) が用いられる。目的とする密度 $\pi(x) = p(x)/K$ から標本を発生するとし、ここで $x \in \mathbb{R}^d$, $p(x)$ は正規化されていない密度であり、 K は (未知の) 正規化定数である。 $g(x)$ をある既知の方法でシミュレートできる密度とし、すべての x に対して $\pi(x) \leq cg(x)$ であるような c が存在するとすると、 $\pi(\cdot)$ から確率変数を得るには

- (1) から確率標本の候補 z を発生させる。
- (2) $(0, 1)$ 区間上の一様分布 $U(0, 1)$ から乱数 u を発生させる。
- (3) $u < \pi(z)/cg(z)$ ならば z を確率標本として受容し、そうでなければ棄却して(1)に戻る。
- (4) (3)で得られた z を用いて $f(\theta|z)$ から確率標本 θ^* を発生させる。

とする。(1)から(3)のステップで受容された z の値は $\pi(\cdot)$ からの変量になることが証明されている (Ripley, 1987)。この方法が有効であるためには、 c を十分注意して選ぶ必要がある。この方法は $g(\cdot)$ の $p(\cdot)$ に対する近似がよくなければ受容率が低くなり、効率の悪い乱数発生法になる。

SIR: Sampling/Importance Resampling 法

もう一つ公表されているデータを用いて統計的推測を簡単に行うために、欠測値を埋める方法として Rubin (1987) は *Sampling/Importance Resampling* (SIR) 法を提案している。まず

²¹⁾ w_i の計算で、 $g(\theta)$ あるいは $p(\theta|y)$ は、任意のスケールファクターを掛けても結果に影響を与えない。しかし、正規化定数は事後オッズ解析に重要であるので、事後オッズを計算するとき注意しなければならない。

$g(Z|y)$ から目標とする m 個の乱数よりもずっと多い M 個の乱数を発生させる。ウェイトに従ってリサンプリングし、目的の m 個の乱数を得る方法である。未知の母数を θ とし、観測されるデータを y 、観測されない観測値を Z とし、事後分布 $p(Z|y)$ から欠測値 Z を発生させる方法である。まず、事後分布 $p(\theta, Z|y)$ に対してよく近似しており、標本を発生させやすい密度関数

$$g(\theta, Z|y) = g(\theta|y)g(Z|\theta, y)$$

を選ぶ。 $g(\theta|y)$ と $g(Z|\theta, y)$ はそれぞれ $p(\theta|y)$ と $p(Z|\theta, y)$ を近似した確率標本を発生しやすい密度関数である。そして

- (1) $g(\theta, Z|y)$ から M 個の標本 (θ_i, z_i) , ($i=1, 2, \dots, M$) を発生する。
- (2) ウェイト w_i を次のように求める。

$$w_j = w(\theta_j, z_j|y) \propto \frac{p(y, z_j|\theta_j)p(\theta_j)}{g(\theta_j, z_j|y)}, j=1, 2, \dots, M$$

- (3) (z_1, z_2, \dots, z_M) の中から $m (< M)$ 個の Z を確率 $p(Z=z_j) = w_j / \sum_{i=1}^M w_i$ で発生させる。

w_j の分布が非常にゆがんでいる場合は SIR による近似はきわめて悪いので、 g の取り方に注意する必要がある。

例：ロジットモデル (続き)

再び Koop and Poirier (1993) のロジットモデルの例に戻ると、ロジットモデルに対する事後密度から直接にサンプリングするのは簡単ではない。結果的に、ベイズ分析で $E(P_{ij}|y)$ のような数量を、解析的あるいは単純なモンテカルロ積分を用いて計算することはできない。この場合は、インポートランスサンプリングを用いたモンテカルロ積分が必要であり、この方法を簡単に導入することができる。

弱い条件の下で、ロジット尤度関数はパラメータに関して凹である。少なくとも Koop and Poirier (1993) で用いられたデータセットに対して、 t インポートランス関数は事後密度に極めてうまく近似できる。上記の論文では、最初にロジット事後分布の対数を最適化し、結果として得た事後モードとマイナスのヘッセ行列の逆行列をインポートランス関数の平均と分散として用いている。自由度のパラメータの様々な値に対して実験し、5つの値を選択すると、それは正規分布の裾よりもかなり厚いことを示している。次に、パラメータや P_{ij} の事後平均や事後分散のような様々な統計量を計算している。

これらのステップを遂行するのに必要なコンピュータプログラムは、非常に単純である。最初のステップでは、ロジット事後密度の対数が最適化され、最適値とヘッセ行列が計算される。もし事前密度がフラットなら、このステップは MLE の計算とその分散の推定と同じである。この最初のステップは、インポートランス関数のパラメータを与える。2番目のステップは、インポートランス関数からサンプリングし、 P_{ij} のような特性値を各サンプリング毎に計算する。次に、そのサンプリングを以前に g^* に対して与えた単純な公式を用いて平均する。

5.3 Poly- t 密度

ベイジアン計量経済学における計算手法を説明する場合に、poly t 密度についても多少ふれる必要があるだろう。Poly t 密度は t カーネル (密度のカーネルは積分定数を無視した密度である) の積あるいは比であり、ベイジアン計量経済分析、特に同時方程式モデルあるいは操作変数モデルでは頻繁に生じる。線形回帰モデルでは、そのような密度は異分散が存在する場合に生じる。これらはしばしば解析的なモーメントをもたないので、数値手法が通常必要とされる。この方法を poly t 事後密度の性質を評価するのに用いることができるけれども、Bayesian

Regression Package (BRP) として知られている特別なプログラムも利用できる (Richard and Tompa, 1980 参照)。このプログラムを詳細に記述することはしない。Poly t 密度で研究しようとする有効な数値積分法はないことを注意すれば十分である。

6. マルコフ連鎖モンテカルロ法

マルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo: MCMC) は、実際の統計モデルの作成にとって大きな展望をもたらした。最近まで、多くの応用問題における分析対象の複雑性と構造を正確に認識する困難性から、そのための特別な方法と特定分析目的のソフトウェアが必要であった。現在では、MCMC 法を利用することによって、多くの複雑な問題を汎用的なソフトウェアを用いて分析することができる統一的な環境が整えられつつある。

MCMC は基本的には、**マルコフ連鎖を用いたモンテカルロ積分**である。ベイズ分析では、場合によっては古典的な分析でさえも、モデルのパラメータについて推測や予測を行うために、かなり高次元の確率分布を積分する必要がある。ベイズ的な方法では、データ所与のもとでのパラメータの事後分布を積分する必要がある。古典的方法はパラメータの値が与えられたもとでの観測値の分布を積分する必要がある。**モンテカルロ積分**は必要な分布から標本を抽出し、期待値に近似するために標本平均を計算する。**マルコフ連鎖モンテカルロ法**は、これらの標本を長期間に巧妙に作られたマルコフ連鎖を作動させることによって抽出する。これらの連鎖を作成する方法は多数あるが、Gibbs サンプラーを含むこれらの方法はすべて、Metropolis et al. (1953) と Hasting (1970) の一般的な枠組みの中の特別なクラスになる。

1990 年代はベイズ的な方法を応用する優れた実証経済分析が表れ始めた年代である。これらの応用のほとんどは事後分布をシミュレートするのに、マルコフ連鎖モンテカルロ法を用いている。シミュレーション・アルゴリズムは、その基本的な形において非常に簡単であるので、多くのベイズ型応用研究ではだんだん標準的な方法になりつつある。さらに、この方法には長い歴史があり、基本的な理論は確立されている。もともとは統計物理の分野で始められ、空間統計学やイメージ解析の分野で長いこと使用されていたが、最近数年の間に MCMC 法はベイズ統計学に大きな影響を与え、古典的な統計分析の分野でもいくつかの応用が見られる。最近の研究では多くの応用、手法の豊富さ、そして理論的な精緻さにおいてかなりの成果が加えられた。

基本的には 2 種類の MCMC 法がある。**ギブス・サンプラー (Gibbs Sampler)** は、完全な条件付き分布 $p(\theta_i | \theta_{j \neq i}, y)$, $i=1, \dots, k$ の集合から順々に標本をとり、かなり緩い条件の下でその安定分布として結合事後分布 $p(\theta|y)$ をもつマルコフ連鎖を生成する。このアルゴリズムは Geman and Geman (1984) によって名付けられ、Gelfand and Smith (1990) は広範囲のベイズ推定問題に適用できることを示した。最近の優れた入門的な解説は Casella and George (1992) を参照するとよい。

もう一つの方法は、完全な条件付き分布からシミュレートするのが難しい時に適用される。この場合、別々の異なるマルコフ連鎖をシミュレートする代わりに、ある別の安定分布を持っているが、それを修正してその安定分布としての事後分布を持つ新しいマルコフ連鎖が生成できるようにする。これは**メトロポリス-ヘイスティングス・アルゴリズム (Metropolis-Hastings Algorithm)** によって達成される。以前に定めた各パラメータに対する**候補分布**から標本をとり、その後受容-棄却ステップを用いる。重要な特徴はこの方法は標本をとる連鎖で正規化されていない事後密度 $p(\theta|y) = L(y|\theta)\pi(\theta)$ しか含まないということである。よく知られた参考文献は、Metropolis et al. (1953) と Hastings (1970) である。それに Chib and Greenberg (1995b) には、優れた入門的な解説がある。

本章では最初に、重要な導入部分であるベイズ推定についてまとめた後にマルコフ連鎖について Chib and Greenberg (1995b) に沿って述べる。次に、Metropolis-Hastings アルゴリズムで与えられる MCMC の一般型を説明した後、ギブス・サンプラーがその特別の場合であることを述べる。応用計量分析でよく用いられているギブス・サンプラーについて述べ、その計算方法と収束判断の方法について簡単にまとめる。最後に、データ拡大アルゴリズムについて述べることにする。

マルコフ連鎖モンテカルロ法に関して、いろいろな面から取り上げた本が最近出版されている。たとえば、Carlin, B. P. and Louis, T. A. (1996) ではデータ解析の観点からベイズ法と経験ベイズ法がまとめられており、Gilks, Richardson, and Spiegelhalter (1996) は 25 編の興味ある論文が基礎から応用まで網羅されている、それに Gamerman (1997) にマルコフ連鎖モンテカルロに属する方法が詳細にまとめて述べられている。Kass, et al (1997) には、この問題に関する興味あるいろいろな示唆が述べられている。日本語で書かれた文献は少ないが、大森 (1996) にはマルコフ連鎖モンテカルロ法の紹介とその収束診断基準が要領よくまとめられている。伊庭 (1997) には統計学の応用が参考になり、繁枘 (1995) の本の付録にこの方法の説明がある。

6.1 ベイズ推定

現在での MCMC 法のほとんどの応用はベイズ推定を行うために始まった。ベイズ分析での観点では、統計モデルの観測可能な変数とパラメータの間には何の基本的な差異はなく、すべて確率変数と考えられている。 y を観測データ、 θ はモデルパラメータと欠測データを表す。推測の段階では、すべての確率変量に関して事前分布 $p(\theta)$ と尤度 $L(y|\theta)$ の 2 つの部分からなる同時分布 $p(y, \theta)$ をセットアップする。 $p(\theta)$ と $L(y|\theta)$ が決まれば、次の確率モデル (*full probability model*) が得られる。

$$p(y, \theta) = L(y|\theta)p(\theta)$$

データ y を観測すると、ベイズの定理を用いて y の条件のもとでの θ の分布を決定する。

$$p(\theta|y) = \frac{p(\theta)L(y|\theta)}{\int p(\theta)L(y|\theta)d\theta}$$

これは θ の事後分布と呼ばれ、すべてのベイズ推定の目的である。

事後分布のすべての特性値、モーメント、四分位値、最高事後密度区域等が、ベイズ推定にとって重要である。これらの統計量は θ の関数の分布全体ではなく、事後期待値によって表すことができる。関数 $f(\theta)$ の事後期待値は

$$E[f(\theta|y)] = \frac{\int f(\theta)p(\theta)L(y|\theta)d\theta}{\int p(\theta)L(y|\theta)d\theta}$$

で得られる。この式の積分は最近まで、特に高次元の場合にはベイズ推定における実際的な困難のほとんどの原因であった。ほとんどの応用分析では、 $E[f(\theta|y)]$ の解析的な評価は不可能である。

期待値の計算

x は k 個の要素からなる確率変数ベクトルで $\pi(\cdot)$ の分布をもつ。ベイズ的な応用分析では、 x はモデルパラメータと欠測データからなり、古典的な応用分析ではデータあるいはランダム効果である。ベイジアンでは $\pi(\cdot)$ は事後分布になり、古典派では尤度になる。どちらも、目的はある関数 $f(\cdot)$ の期待値を計算することである。

$$E[f(x)] = \frac{\int f(x)\pi(x)dx}{\int \pi(x)dx}$$

ここで x の分布は正規化定数まで既知である可能性を認めることにする。たとえば、ベイジアン推定では $p(\theta|y) \propto p(\theta)L(y|\theta)$ は知っているが、正規化定数 $\int p(\theta)L(y|\theta)d\theta$ を簡単に評価することができない。

モンテカルロ積分

モンテカルロ積分は、目標とする密度 $\pi(\cdot)$ から標本 $\{x_t, t=1, \dots, n\}$ を抽出することによって $E[f(x)]$ を評価し、次に

$$E[f(x)] \approx \frac{1}{n} \sum_{t=1}^n f(x_t)$$

で近似する。 $f(x)$ の母集団平均は標本平均によって推定される。標本 $\{x_t\}$ が独立なら、大数の法則によって近似は標本サイズ n を増やすだけ正確になる。ここで n は分析者がコントロールすることができ、固定されたデータサンプルのサイズではない。

一般に、 $\pi(\cdot)$ は標準的ではないので、標本 $\{x_t\}$ は $\pi(\cdot)$ から独立に抽出されることはない。しかし、 $\{x_t\}$ は必ずしも独立である必要ではない。 $\{x_t\}$ は任意のプロセスによって生成することができる。大胆に言えば、 $\pi(\cdot)$ のサポートを通じて正しい割合で標本を取ればよい。これを行う一つの方法は、安定分布として $\pi(\cdot)$ をもつマルコフ連鎖によって行う。これが**マルコフ連鎖モンテカルロ**である。

6.2 マルコフ連鎖

各時間 $t \geq 0$ で一連の確率変数 $\{x_0, x_1, x_2, \dots\}$ を生成すると仮定すると、次の状態 x_{t+1} は連鎖の現在の状態 x_t にのみ依存する分布 $P(x_{t+1}, x_t)$ からとられる。これは、 x_t 所与のもとでの次の状態 x_{t+1} は、過去の連鎖 $\{x_0, x_1, \dots, x_{t-1}\}$ にはこれ以上依存しない。この系列を**マルコフ連鎖**と呼び、 $P(\cdot, \cdot)$ は連鎖の**推移核**と呼ぶ。連鎖は時間同時性、すなわち $P(\cdot, \cdot)$ は t に依存しないと仮定する。一般的には遷移核 $P(x, A)$ は条件付き分布関数で、 x から集合 A 内の一点へ移動する確率を表している。分布関数であるから、 $P(x, \mathbb{R}^d) = 1$ で、ここで連鎖は点 x から x へ移動する場合もある、すなわち $P(x, \{x\})$ は必ずしもゼロにならない。

初期状態 x_0 がどのように x_t に影響を与えるのかを見るために、 x_0 所与のもとでの x_t の分布を $P^{(t)}(x_t, x_0)$ と表す。ここで中間の値 $\{x_1, x_2, \dots, x_{t-1}\}$ は与えられていないので、 x_t は直接 x_0 に依存する。正則条件のもとで、連鎖は徐々にその初期状態を忘れ、 $P^{(t)}(\cdot, x_0)$ は究極的には一意な**安定分布** (あるいは**不変分布**) に収束し、これは t あるいは x_0 に依存しない。当面、安定分布を $\phi(\cdot)$ と表す。すると、 t を増やすにつれ、標本点 $\{x_t\}$ は $\phi(\cdot)$ からの相関のある標本のよう増加しているように見える。

連続状態空間に対するマルコフ連鎖理論の方法は、遷移核 $P(x, A)$ から出発する。遷移核は条件付き分布関数で、 x から集合 A 内の一点へ移動する確率を表している。分布関数であるから $P(x, A)$ で、ここで連鎖は点 x から x へ移動する場合もある、すなわち $P(x, \{x\})$ は必ずしもゼロにならない。

マルコフ連鎖理論の主要な問題は、安定分布 π^* が存在する条件と遷移核のイタレーションが安定分布へ収束する条件を決めることである。安定分布は、次の条件を満足する。

$$\pi^*(dy) = \int P(x, dy)\pi(x)dx \tag{1}$$

問題は適切な $P(x, dy)$ を見付けることである。関数 $p(x, y)$ に対して遷移核は次のように表される。

$$P(x, dy) = p(x, y)dy + r(x)\delta_x(dy) \quad (2)$$

ここで、もし $x \in dy$ なら $p(x, x) = 0$, $\delta_x(dy) = 1$, その他はゼロである。そして $r(x) = 1 - \int p(x, y)dy$ は連鎖が x にとどまる確率である。 $r(x) \neq 0$ の可能性から、 y に関して $p(x, y)$ の積分が必ずしも 1 にならないことは明らかである。

さて、関数 $p(x, y)$ が反転条件 (あるいは詳細均衡)

$$\pi(x)p(x, y) = \pi(y)p(y, x) \quad (3)$$

を満たせば、 $\pi(\cdot)$ は $P(x, \cdot)$ の不変密度になる。

6.3 Metropolis-Hastings アルゴリズム

メトロポリス・ヘイスティングス・アルゴリズムは、標準的でない分布 $\pi^*(\cdot)$ からサンプルするのに用いることができる強力な MCMC 法である。アルゴリズムから一連の抽出を行う方法は、その安定分布 $\phi(\cdot)$ が関心ある分布 $\pi(\cdot)$ であるようなマルコフ連鎖を生成する方法を見つける必要があるが、このようなマルコフ連鎖を生成するのは非常にやさしい。以下で、その形を Hastings (1970) にしたがって述べる。これは Metropolis et al. (1953) によって最初に提案された方法の一般化である。

前述の受容/棄却法と同様に、いくつかの候補を生成できる密度があると仮定する。マルコフ連鎖を用いていることから、密度がプロセスの現在の状態に依存することを認める。したがって、候補生成密度 (candidate generating density あるいは proposal density) は $q(x, y)$ と表され、ここで $\int q(x, y)dy = 1$ である。この密度はプロセスが x 点にあるとき、密度は y を $q(x, y)$ から生成されることを意味している。もし $q(x, y)$ 自身が反転条件を満たせば探索は終了するが、ほとんどそういうことはない。たとえば、ある x, y について

$$\pi(x)q(x, y) > \pi(y)q(y, x) \quad (4)$$

となる場合がある。この場合、大ざっぱには、プロセスは x から y へ頻繁に移動し、 y から x へは滅多に動かない。この状態を修正する簡単な方法は、移動が起こる確率 $\alpha(x, y) < 1$ を導入することによって x から y へ移動する数を減らす。ここで $\alpha(x, y)$ を移動確率と呼ぶ。移動が起こらなければ、プロセスは再び x を目的分布からの値として戻す。そこで x から y ($y \neq x$) への移行は

$$p_{MH}(x, y) \equiv q(x, y)\alpha(x, y), \quad x \neq y$$

に従って起こる。ここで $\alpha(x, y)$ も決められる。

再び不等式 (4) を考えると、これは y から x へはそれほど起こらないことを表している。したがって、 $\alpha(x, y)$ をできるだけ大きくするようにし、これは確率であるから上限は 1 になる。しかし、現在は移動確率 $\alpha(x, y)$ は $p_{MH}(x, y)$ が反転条件を満足するように決められる。そのとき次のようになるためである。

$$\begin{aligned} \pi(x)q(x, y)\alpha(x, y) &= \pi(y)q(y, x)\alpha(y, x) \\ &= \pi(y)q(y, x) \end{aligned}$$

ここから $\alpha(x, y) = \pi(y)q(y, x) / \pi(x)q(x, y)$ であることが分かる。もちろん不等号が逆の時は $\alpha(x, y) = 1$ である。確率 $\alpha(x, y)$ と $\alpha(y, x)$ は前の不等式の両辺をバランスするように、言い換

えると $p_{MH}(x, y)$ が反転条件を満足することを保証するために決められる。そこで、 $p_{MH}(x, y)$ が反転可能であるためには、移動確率は次のように決められる。

$$\alpha(x, y) = \begin{cases} \min \left[\frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}, 1 \right] & \pi(x)q(x, y) > 0 \text{ のとき} \\ 1 & \text{その他} \end{cases} \quad (5)$$

M-H 連鎖に対する遷移核の定義を行うために、プロセスが x にとどまるゼロでない確率を考える。先に定義したように、この確率は

$$r(x) = 1 - \int_{\mathcal{Y}} q(x, y) \alpha(x, y) dy$$

となる。したがって、 $p_{MH}(x, dy)$ と表される M-H 連鎖の遷移核は

$$P_{MH}(x, dy) = q(x, y) \alpha(x, y) dy + \left[1 - \int_{\mathcal{Y}} q(x, y) \alpha(x, y) dy \right] \delta_x(dy)$$

で与えられる。

Metropolis-Hasting アルゴリズムでは、各期 t で次の状態 x_{t+1} は提案された分布 $P(\cdot, x_t)$ から候補点 y を最初に抽出する。提案された分布は現在の点 x_t に依存していても良い。たとえば、 $P(\cdot, x_t)$ は多変量正規分布で、平均が x 、一定の共分散行列である場合もある。候補点 y は、次の確率 $\alpha(x_t, y)$ で受容される。最後のサンプルで x の値を得たとすると、次の値は候補生成密度 $q(x, y)$ から y をサンプルすることによって生成される。このようにして得られた y は、確率 $\alpha(x, y)$ で受容される。候補点を受容されると、次の状態は $x_{t+1} = y$ となる。確率 $1 - \alpha(x, y)$ で候補が棄却されると、連鎖は動かずサンプルした値は現在の値 $x_{t+1} = x_t$ となる。

このマルコフ連鎖の状態遷移を決める移動確率は、反転条件を満たしてマルコフ連鎖は時間に可逆 (reversible) である。このことから

$$\int \pi(x_t) q(x_{t+1}, x_t) dx_t = \pi(x_{t+1})$$

が言え、 $\pi(x)$ はこのマルコフ連鎖の定常分布になる。

もし $x^{(0)} \sim \pi(x)$ なら、この式より各 $x^{(t)}$ 、 $t \geq 1$ も $\pi(x)$ に従うことになるが、 $x^{(0)}$ をそのように選ぶことはできない。実際には、十分に長い、たとえば m 回のイタレーションの後、点 $x^{(m+1)}$ 、 \dots 、 $x^{(n)}$ まで発生させ、最初の m 個のサンプルを捨ててマルコフ連鎖からの結果を用いて、期待値 $E[f(x)]$ を推定する。ここで x は分布 $\phi(\cdot)$ をもつ。初期のサンプルを捨てることは "burn-in" と呼ばれる。この burn-in 標本は通常計算から除外され、次の推定量を得る。

$$\bar{f} = \frac{1}{n-m} \sum_{t=m+1}^n f(x_t) \quad (6)$$

これはエルゴートの平均と呼ばれる。求める分布への収束は、エルゴード定理により保証される。一般にマルコフ連鎖が proper な定常分布を持ち、既約 (irreducible) で非周期的 (aperiodic) なら、定常分布は一意で、 $t \rightarrow \infty$ の時の $x^{(t)}$ の極限分布が $\pi(x)$ に収束する。

いくつかの重要な点がある。第 1 に、 $\alpha(x, y)$ に計算には $\pi(\cdot)$ の正規化定数の知識を必要としない。第 2 に、もし候補生成密度が対称なら、すなわち $q(x, y) = q(y, x)$ なら、受容確率は $\pi(y)/\pi(x)$ に縮小し、これは Metropolis et al. (1953) の本来の公式である。最後に、Gibbs サ

ンプラーはMHアルゴリズムの特別な場合であることが示すことができる (Chib and Greenberg, 1995b).

6.4 Gibbs サンプリング法

5章で述べた手法は、様々な計量経済モデルを分析することができるが、多くの場合事後密度から直接サンプリングしたり、あるいは良いインポートランス関数を見つけるのが不可能な場合があり、その場合は上で述べたモンテカルロ法を直接適用することができない。単純なモンテカルロ積分を事後密度からサンプリングし、これらのサンプルから事後密度の任意の特性値の期待値が計算できることを前に述べた。Gibbs サンプリング法は事後分布からのサンプリングに収束する乱数を生成するので、この性質のためにこの方法がよく使われるようになった。これらの乱数から、モンテカルロ積分が利用できない場合における事後密度の特性値を簡単に計算することができる²²⁾。単一要素だけからなる Metropolis-Hastings の特別な場合が **Gibbs サンプリング**である。Gibbs サンプリングはその名称を Geman and Geman (1984) によって与えられ、格子状の Gibbs 分布を解析するのに用いられたが、その応用可能性は Gibbs 分布に限られない。そこで Gibbs サンプリングは実際は誤解しやすい。同じ方法はすでに統計物理の分野で使われており、そこでは **heat bath algorithm** として知られている。言うまでもなく、Geman and Geman (1984) の仕事は Gelfand and Smith (1990) と Gelfand et al. (1990) の研究を通じて MCMC の統計学の主要な分析道具への道を拓いた。今日まで、MCMC のほとんどの統計的な応用は Gibbs サンプリングが用いられている。

最近、応用計量経済分析を行う人々の間に Gibbs サンプリング法に対する関心が生じてきた。ギブス法は、次のような分野の分析に用いられている。多項目反応データ (Albert and Chib, 1993a)；トービットモデル (Chib, 1992)；多対プロビットモデル (McCulloch and Rossi, 1994)；ARMA(p, q) 誤差を持つ回帰モデル (Chib, 1993；Chib and Greenberg, 1992)；マルコフスイッチング・モデル (Albert and Chib, 1993b；McCulloch and Tsay, 1994a)；ストカステック・フロンティア・モデル (Broeck, Koop, Osiewalski and Steel, 1994)；ストカステック・ボラティリティ・モデル (Jacquier, Polson and Rossi, 1993)；多重変化点モデル (Chib, 1998)，そして非共役事前分布をもつ時系列モデル (Geweke, 1992a) である。この分野での最近の研究の発展から判断すると、ギブス法は応用ベイズ計量経済分析には重要な手法になっていることが分かる。

Gibbs サンプリング法の主要な仮定は、たとえ同時分布からサンプリングするのが難しくとも、いくつかの条件付分布からは簡単にサンプリングできることである。たとえば、 A と B が確率変数であるとし、 $p(A, B)$ からサンプリングしたい。たとえ $p(A, B)$ から直接にサンプルをとることができなくても、 $p(A|B)$ と $p(B|A)$ からサンプリングできれば、Gibbs サンプリングを使って $p(A, B)$ から極めて簡単にサンプリングができる。特に、条件付分布から次々とサンプルをとることによって、同時密度からの標本に収束することができる。条件なしの密度からよりも、条件付密度からサンプリングする方がずっと簡単であるから、ギブス・サンプリングは様々な計量経済学上の問題に対して役に立つことを示した。

もっと正確には、 θ を $(\theta_1, \dots, \theta_k)$ のように分解されるパラメータ・ベクトルとする。ギブス・サンプリングが役に立つためには、 $j=1, \dots, k$ に対して $p(\theta_j | \theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_k)$ から簡単にサンプリングできるような分割を選ばなければならない。Gibbs サンプリングは、各サンプルが他の密度からの以前のサンプルについての条件付きである密度から連続してサンプリングを行

²²⁾ Gibbs サンプリングはこのような場合に適しているが、他の似たような方法も存在する。関心がある人は、Tierney (1991), Tanner (1991), Gibbs サンプリングと関連手法に対する議論は、Tanner and Wong (1987) あるいは Gelfand and Smith (1990) を参照すると良い。

う。すなわち、 θ_j^i を $j=1, \dots, k, i=1, \dots, s$ に対する $p(\theta_j | \dots, \theta_{j-1}^{i-1}, \theta_{j+1}^{i-1}, \dots, \theta_k^{i-1})$ からのサンプルとし、ここで θ^0 は任意の初期値である。 k はステップ数(抽出される条件付密度の数)を示し、 s はパスの数を示す。もし θ^i を i 番目のパス、 $\theta^i = (\theta_1^i, \dots, \theta_k^i)$ 、でとられた1組のサンプルとすると、弱い条件の下で θ^i は $p(\theta)$ からのサンプルに収束する。さらに、次を示すこともできる。

$$g^* = \frac{1}{s} \sum_{i=1}^s g(\theta^i) \rightarrow E(g(\theta))$$

ここで例示のために、2変量の場合を考える。初期値 $\theta^0 = (\theta_1^0, \theta_2^0)$ を与えると、ギブス・サンプラーは1つのサンプル θ_1^1 を $p(\theta_1 | \theta_2^0)$ からとり、1つのサンプル θ_2^1 を $p(\theta_2 | \theta_1^1)$ からとり、1つのサンプルを $p(\theta_1 | \theta_2^1)$ からとり、等々と繰り返す。何回かのパスの後、この方法によって $p(\theta_1, \theta_2)$ から連続してサンプルをとり、 $E(g(\theta))$ 自身と $g(\cdot)$ を計算するのに使うことができる。

一般的には、 $\pi(\theta), \theta \in S \subseteq \mathbb{R}^p$ を(正規化されていない)目標密度とする。 θ を $\theta_1, \theta_2, \dots, \theta_d$ に何らかの分解し、 k 番目のブロックの完全条件付き密度を $\pi(\theta_k | \theta_1, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_d)$ と表す。すると、Gibbs サンプリング・アルゴリズムは次の繰り返しによって定義される。

1. 初期値 $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})$ を与え、 $i=0$ とおく
2. 次のようにシミュレートする

$$\begin{aligned} \theta_1^{(i+1)} & \text{ を } \pi(\theta_1 | \theta_2^{(i)}, \theta_3^{(i)}, \dots, \theta_d^{(i)}) \text{ から} \\ \theta_2^{(i+1)} & \text{ を } \pi(\theta_2 | \theta_1^{(i+1)}, \theta_3^{(i)}, \dots, \theta_d^{(i)}) \text{ から} \\ \theta_3^{(i+1)} & \text{ を } \pi(\theta_3 | \theta_1^{(i+1)}, \theta_2^{(i+1)}, \dots, \theta_4^{(i)}, \dots, \theta_d^{(i)}) \text{ から} \\ & \vdots \\ \theta_d^{(i+1)} & \text{ を } \pi(\theta_d | \theta_1^{(i+1)}, \theta_2^{(i+1)}, \dots, \theta_{d-1}^{(i+1)}) \text{ から} \end{aligned}$$

3. $i=i+1$ とおいてステップ2に戻る

このアルゴリズムはマルコフ連鎖の次の項 $\theta^{(i+1)}$ を各完全条件付き密度をシミュレートすることによって与え、ここで条件付きの要素は一つのサイクル中に改訂される。

次に、Gibbs サンプリングを実行する際に生じるいくつかの問題に移ろう。最初に、ブロックを計画するとき高い相関がある要素は一緒にまとめるべきである。そうしないと、マルコフ連鎖はゆっくりとしか減衰しない自己相関を示す傾向があり、目標密度へゆっくりと収束することになる。第2に、扱いやすい完全条件付き構造は、しばしば x の定義に潜在データあるいは欠測データを導入することによって得られる。サンプラーに変数を追加するという考えはデータ拡大として知られており、Tanner and Wong (1987) で導入され、いくつかの応用例がある。最後に、完全条件付き密度のいくつかが伝統的な方法(棄却サンプリング法や既知の発生器)によってサンプルすることが難しい場合、その密度はMHアルゴリズムによって(Müller, 1991)か、あるいは独立サンプルを生成することによってサンプルすることができる(Gilks and Wild, 1992)。

各パスの最後で、ギブスサンプラーは $\theta_1, \dots, \theta_k$ のそれぞれに対してサンプルを得る。パス毎の平均を、任意の $g(\cdot)$ に対する $E(g(\theta) | y)$ を推定するのに使うことができる。しかし、2つの問題がモンテカルロ積分を適用できないGibbs サンプリング手法を使う際に生じる。それは初期条件とパス毎の系列相関である。理論的には、ギブスサンプラーから得るサンプルは、パスの数が多きときだけは事後分布からのサンプルと等しくなる。もし θ^0 が不適切に選ばれると、初期のサンプルは後のサンプルと大きく異なる場合がある。このために、多くの分析ではギブスサンプラー中の m 個の初期のパスを棄てており、これによって θ^0 の選択への依存を減らす。

単純なモンテカルロ積分とは異なり、ギブスサンプラーは事後密度からの独立な標本を生成せず、したがって、サンプルがパス毎に系列相関があることを簡単に示すことができる。この特性から、研究者によっては系列相関を断ち切る意味から r 回目のパス毎のサンプルを棄てるか、 r 回目のパス毎にギブスサンプラーを再始動させている。実証的な問題では様々な方策を試みるか、 θ^0, m, r の異なる選択をし、結果がこれらに敏感かどうかを見ることが重要である。もちろん、最適な Gibbs サンプリング・アルゴリズムは問題毎に異なる。しかし、ギブスサンプラーから 1 つの長いランを得た方がよいか、あるいは再出発した方がよいか？ という問題に関心が持たれている (Tanner, 1991; Zeger and Karim, 1991; Raftery and Lewis, 1991; それに Gelman and Rubin, 1992; を参照)。

モンテカルロ法が正規近似あるいはラプラス近似より優る 1 つの利点は、この方法には診断法があり、近似の精度を反復数を選択することによりコントロールすることができることにある。Geweke (1992a) は、ギブス・サンプラーに対して同じ種類の診断を提案している。すなわち、もし $E(g(\theta)|y)$ の推定値を

$$g^* = \frac{1}{s^*} \sum_{i \in I} g(\theta^i)$$

とし、ここで I は推定に含まれる一組のパスを示す (I はすべてのパス、あるいは最初の m を除いたすべて、などでも良い)、そして s^* は含まれるサンプル数 (s あるいは $s-m$) とすると、1 組の $\{g(\theta^i)\} (i \in I)$ は単変量確率過程になる。漸近理論を用いて、Geweke は g^* は平均 $E(g(h)|y)$ 、漸近分散 $s^{*-1}S_{(0)}$ の漸近正規密度にしたがうことを示した。ここで $S_{(0)}$ は周波数がゼロのところで評価した $\{g(\theta^i)\}$ のスペクトル密度である。この統計量は、 $[s^{*-1}S_{(0)}]^{1/2}$ は g^* に対する NSE として用いることができ、数値標準誤差の計算はギブスサンプラーを採用すると可能になることを示した。 $S_{(0)}$ の推定値を得る多くの方法があり、これらは標準的な時系列のテキストで述べられている。さらに、RNE はよく知られた方法で計算することができる。

Geweke (1992b) はさらに別の収束診断法を作成した。ギブスサンプラーはパスの数が大きくなったときにだけ、事後分布からのサンプルを生成する。したがって、初期のパスを後のパスと比較すると収束の失敗が明らかになる。 $E(g(\theta)|y)$ の 2 つの推定値が作成され、次のように 1 つは初期のパスに基づくものであり、もう 1 つは後のパスに基づくものである。

$$g_A = \frac{1}{s_A} \sum_{i=1}^{s_A} g(\theta^i), \quad g_B = \frac{1}{s_B} \sum_{i=s_c}^{s^*} g(\theta^i)$$

ここで $s_c = s^* - s_B + 1$ 。もし 2 つの推定値が大きく変われば、これは収束しなかったことを示す。より詳しくは、 NSE_A と NSE_B を前のパラグラフでのように計算した 2 つの推定値に対する NSE とするとして、Geweke は次を示した。

$$\frac{(g_A - g_B)}{(NSE_A^2 + NSE_B^2)^{1/2}} \rightarrow N(0, 1)$$

もし $[g(\theta)]$ が定常なら、 s_A/s^* と s_B/s^* は固定されており、 $(s_A + s_B)/s^* < 1$ である。この応用では、Geweke は $s_A = 0.1s^*$ と $s_B = 0.5s^*$ とおいた。彼は上の方程式の左辺を収束診断とし、これを収束を確かめるのに用いることを薦めている。

本節では、Gibbs サンプリング法が基本的なモンテカルロ法の強力な拡張になっていることを述べた。ギブスサンプラーは同時事後密度がそれ自体扱い難いが、パラメータ空間を適切に分割した条件付分布から抽出しやすい場合には適している。ギブスサンプラーを利用した論文

数の最近の急増ぶりから判断すると、計量経済学の多くの問題は、そのような条件付の密度によってより簡単に扱えるようになるように見える。

例：マルコフ型スイッチング回帰モデル

次の McCulloch and Tsay (1994) で用いられたマルコフ型スイッチング回帰モデルの単純化したものを考える。

$$\begin{aligned} y_t &= X_t' \beta_1 + \varepsilon_{1t} & s_t = 1 \text{ のとき} \\ y_t &= X_t' \beta_2 + \varepsilon_{2t} & s_t = 2 \text{ のとき} \end{aligned}$$

ここで ε_{it} は、 $iid. N(0, \sigma_i^2)$ の確率変数 ($i=1, 2, t=1, \dots, T$) である。任意の時間 t で、これらの回帰モデルのどちらか一方が成り立ち、この状態間のスイッチはマルコフ過程にしたがう。

$$\begin{aligned} p(s_t = 2 | s_{t-1} = 1) &= p_1, \\ p(s_t = 2 | s_{t-1} = 2) &= p_2. \end{aligned}$$

いまパラメータ $\theta = (\beta_1, \beta_2, \sigma_1^2, \sigma_2^2, p_1, p_2)$ と状態ベクトル $s = (s_1, \dots, s_T)$ に関心がある。このモデルの尤度関数は複雑である。実際、たとえ非報知事前密度を仮定しても、事後密度から直接にサンプリングするのは難しい。しかし、状態ベクトルを知っていれば、モデルは2つの単純な回帰を含んでいる。そして、パラメータの値とデータが既知であれば、状態について推測するのは簡単である。このことは、このモデルに対してギブス・サンプラーをセットアップできることを示唆している。

McCulloch and Tsay (1992a) は、ギブスサンプラーが次の条件付分布でセットアップできることを示した。 $i=1, 2$ に対して $p(\beta_i | y, s, \sigma_i^2)$, $p(\sigma_i | y, \beta_i)$, $p(p_i | s)$; そして $t=1, \dots, T$ に対して $p(s_t | y, s_{(-t)}, \beta_1, \beta_2, \sigma_1^2, \sigma_2^2)$, ここで $s_{(-t)} = (s_1, \dots, s_{t-1}, s_{t+1}, \dots, s_T)$ とする。さらに、これらの条件付サンプルはそれぞれ簡単に抽出できる。特に、 $p(\beta_i | y, s, \sigma_i^2)$ が正規密度、 $p(\sigma_i | y, \beta_i)$ は逆カイ2乗密度、 $p(p_i | s)$ はベータ密度、そして $p(s_t | y, s_{(-t)}, \beta_1, \beta_2, \sigma_1^2, \sigma_2^2)$ は状態ベクトルのマルコフ性を利用することによって、簡単に操作できる密度である。証明と正確な密度関数の詳細は、McCulloch and Tsay を参照すると良い。

ギブス・サンプラーは、各サンプルが以前のサンプルの条件付きであるときに、これらの条件付分布から連続してサンプリングする。McCulloch and Tsay (1994) はこのタイプのモデルを用いて米国のGNPと失業率を調べ、McCulloch and Tsay (1994) は同様のモデルを用いて米国の工業生産における単位根の存在を推定している。この種のモデルは最尤法では分析するのが非常に難しいので、ギブスサンプラーを用いたベイズ推定が極めて自然である。

例：打ち切りデータのある線形回帰モデル

次の線形回帰モデルを考える。

$$w_t = x_t' \beta + \varepsilon_t$$

ここで x_t は t 番目の個人に対する k 次元の外生変数ベクトル、そして ε_t は $iid. N(0, \sigma^2)$ 、しかし w_t は観測されない変数である。そのかわり y が観測され、ここで $y_t \leq y^*$ なら $y_t = w_t$ 、その他の場合は $y_t = y^*$ である(従属変数はセンサーされている)。もし従属変数がセンサーされていなければ、モデルは扱いやすい(事後密度は正規誤差を持つ線形回帰モデルに対する事後密度と同じになる)。すなわち、 w_t が既知なら事後密度は操作可能であるが、もしパラメータが既知なら、 w_t についての推測は極めて簡単に行える。これは、連続したサンプルを $p(\beta, \sigma^2 | y, w)$ と $p(w | y, \beta, \sigma^2)$ からとるギブスサンプラーを示している。非報知事前密度、あるいは共役事前密

度を仮定すると、 $p(\beta, \sigma^2|y, w)$ は正規-逆ガンマ密度、 $p(w|y, \beta, \sigma^2)$ は多変量正規密度であることを示すことができる。したがって2つの条件付分布は簡単にサンプルでき、ギブスサンプリング法は拡張された事後密度の分析に採用することができる。より詳細は、Chib (1992) を参照のこと。

6.5 データ拡大 (Data Augmentation) アルゴリズム

Tanner and Wong (1987) によって開発されたデータ拡大 (Data Augmentation) アルゴリズムは、観測できるデータ y に観測できない潜在変数 z を人工的に発生して拡大し、パラメータの事後分布の形状等を評価する方法である。 y と z が与えられれば事後分布 $p(\theta|y, z)$ が簡単になることを仮定しており、この方法は次の imputation (埋め込み) と posterior (事後) ステップからなる。

imputation step 予測分布 $p(z|y)$ に対する現在の近似分布から、 m 個の潜在変数 z の標本 z_1, z_2, \dots, z_m を発生させる。

posterior step 得られた拡大データ (y, z) を用いて、モンテカルロ法により $p(\theta|y)$ を近似する。

imputation ステップにおいて、 $p(z|y)$ は

$$p(z|y) = \int p(z|\theta, y)p(\theta|y)d\theta$$

であり、 $p(\theta|y)$ に関する知識を前提とする。すなわち、埋め込みステップと事後ステップを繰り返すことを前提とする。具体的には

- (1) 第 i 回目の $p(\theta|y)$ の近似を $p^{(i)}(\theta|y)$ とする。 $p^{(i)}(\theta|y)$ から θ^* を発生する。
 - (2) θ^* を所与として $p(z|y)$ から z を発生する。
- (1)と(2)を繰り返すことによって $p(z|y)$ から m 個の z^* を発生する。 $p(\theta|y)$ の近似は次の積分に基づく。

$$p(\theta|y) = \int p(\theta|y, z)p(z|y)dz$$

この積分を m 個の z^* を使ったモンテカルロ法によって近似する。

$$p^{(i+1)}(\theta|y) = \frac{1}{m} \sum_j^m p(\theta|y, z_j)$$

埋め込みステップの $p^{(i)}(\theta|y)$ もこのようにして得られる。このような混合分布に従う乱数を、最初に m 個の分布の一つを無作為に選択し、次に選ばれた $p(\theta|y, z_j)$ から θ を1つ発生する様にして作成する。

この考え方を応用したモデルには、ランダム・ポイントにおける構造変化モデル (Carlin, Gelfand, and Smith, 1992), 打ち切りデータや離散データのモデル (Chib, 1992; Albert and Chib, 1993b), マルコフ・スイッチングモデル (Albert and Chib, 1993b; Chib, 1993; McCulloch and Tsay, 1994), パラメータ制約のあるモデル (Gelfand, Smith, and Lee, 1992), それにセンサー化自己回帰モデル (Zangari and Tsurumi, 1996) 等々が含まれる。

例：センサー化自己回帰モデル

Zangari and Tsurumi (1996) では、センサー化自己回帰誤差モデルのパラメータの事後モーメントと事後確率密度を3つの方法で導いている。ラプラス近似、データ拡大のある Gibbs サンプリング、そしてガウス求積法である。ベイズ点推定量を事後平均として求め、それをシ

ミュレートした EM アルゴリズムと疑似最尤法で求めた最尤推定量と比較した。

これらの方法は、日本の米国への乗用自動車輸出の 1974 年から 1992 年までの年データと四半期データを用いた回帰モデルに適用された。日本の乗用自動車輸出は、1981 年から 1986 年まで輸出自主規制として知られている実質数量割り当ての状態におかれていた。この自主規制は米国の自動車生産が自主規制の上限が無効になるまで 6 年間続いた。この間の時期をセンサーされていたと考え、自動車輸出の需要関数を計測している。彼らの研究の結果、標準的な Tobit 回帰モデルとは異なり、センサーされている観測値がかなり多く、データに強い自己相関が見られる場合 Gibbs サンプラーと結びつけたデータ拡大アルゴリズムは計算上の問題に直面することが報告されている。

センサーされた観測値 $y^c = \{y_i^c\}$ とともに、対応する潜在データ $y^{c*} = \{y_i^{c*}\}$ が利用できるとする。さらに、センサーされていない観測値 $y = y^*$ を $y' = \{y_i'\}$ と表す。結果的に、 $y = \{y', y^c\}$ と $y = \{y_i\}$ である。データ拡大 Gibbs サンプラーの基礎は $p(\theta|y, y^*) = p(\theta|y', y^{c*})$ 、すなわち $p(\theta|y^c, y', y^{c*}) \stackrel{d}{=} p(\theta|y', y^{c*})$ である。 y^{c*} は観測されなくとも、シミュレートできる。センサー化された自己相関のある回帰の変換は、もとの式からその 1 期ラグを引いて得られる。すなわち、 $y_i^c = (1 - \rho L)y_i^*$ と $x_i^c = (1 - \rho L)x_i$ 、ここで L はラグオペレータである。変換したデータでのモデルは、次のように表される。

$$y^0 = x^0 \beta + u, \quad u \sim \mathcal{N}(0, \sigma^2 I_{T-1})$$

事前分布 $p(\beta, \sigma^2)$ は ρ とは独立であるとする、 (β, σ^2, ρ) の事前分布は、 $p(\beta, \sigma^2, \rho) = p(\beta|\sigma^2)p(\sigma^2)p(\rho)$ となる。ここで、特に

$$p(\beta|\sigma^2) \sim \mathcal{N}(\beta_0, \sigma^2 A_0^{-1}), \quad p(\sigma^2) \sim \mathcal{IG}(\nu_0/2, \delta_0/2), \quad p(\rho) \sim \mathcal{N}(\psi_0, \Psi_0^{-1})$$

とする。Gibbs サンプラーでは、次の条件付き分布が必要である： $p(\beta|\sigma^2, \rho, y, y^*)$, $p(\sigma^2|\beta, \rho, y, y^*)$, $p(\rho|\beta, \sigma^2, y, y^*)$, そして $p(y^{c*}|\beta, \sigma^2, \rho, y)$ である。 y^{c*} の条件付きの構造は

$$p(y^{c*}|y, \beta, \rho, \sigma^2) \sim \text{truncated-}\mathcal{N}_{vc}(\eta^c, \Sigma^c)$$

であるから、 β の正規事前分布を尤度関数を結びつけて、次式を得る。

$$p(\beta|y, \rho, \sigma^2, y^c) \sim \mathcal{N}(\beta^*, \sigma^2 A^{*-1})$$

ここで、 $\beta^* = (A_0 + x^0 x^0)^{-1}(A_0 \beta_0 + x^0 y^0)$, $A^* = (A_0 + x^0 x^0)$ である。 β と ρ が与えられたときの σ^2 の尤度はすぐに得られ、仮定された事前分布は

$$p(\sigma^2|\beta) \sim \mathcal{IG}((\nu_0 + k)/2, (\delta_0 + Q_\beta)/2), \quad p(Q_\beta) = (\beta - \beta_0)' A_0 (\beta - \beta_0)$$

を意味する。 σ^2 の条件付き分布は

$$p(\sigma^2|y, \rho, \beta, y^{c*}) \sim \mathcal{IG}((T-1 + \nu_0 + k)/2, (\delta_0 + Q_\beta + d_\beta)/2)$$

ここで、 $d_\beta = (y^0 - x^0 \beta)(y^0 - x^0 \beta)$ である。 $p(\rho|y, \sigma^2, \beta, y^{c*})$ の条件付き事後分布は、次のように導出される。AR(1) モデルでは、誤差 $\epsilon_t = y_t^* - x_t \beta$ は次のように書く

$$\epsilon_t = \rho \epsilon_{t-1} + u_t, \quad t = 2, \dots, T$$

ρ は単変量正規分布に従う。尤度関数と ρ の正規事前密度を結合すると、次の切断された正規事後密度を得る。

$$p(\rho|y, \sigma^2, y^{c*}, \beta) \propto \text{truncated-}\mathcal{N}(\psi, \Psi^{-1})$$

ここで、 $\epsilon = (\epsilon_1, \dots, \epsilon_T)'$, $\psi = \Psi^{-1}(\Psi_0 \phi^0 + \sigma^{-2} M \epsilon)$, $\Psi = (\Psi_0 + \sigma^2 M M')$ である。そして M は $(T \times 1)$ ベクトルで、その第 i 行は ϵ_{t-1} である。

7. 計量経済分析における MCMC シミュレーション法

MCMC 法は、このようにして目標分布 (*target distribution*) から標本 (多重観測値) を生成するシミュレーション技法である。マルコフ過程の推移核は、その極限の不変分布が目標分布になるという性質を持つように定式化され、マルコフ連鎖はコンピュータで生成されるモンテカルロ・シミュレーションを多数回反復する。そして一時的な不安定状態 (*burn-in*) の後、様々な条件の下での出力が目標分布からのサンプルになる。そのような一般的な方法は、Metropolis-Hastings (MH) アルゴリズムとして知られている。このアルゴリズムでは、マルコフ連鎖の次の値は提案密度から生成され、現在の点での密度と比べたときの候補点での密度に従って、受容したり棄却されたりする。MH 法の特別な場合は Gibbs サンプルング・アルゴリズムで、Geman and Geman (1984) で導入され、Tanner and Wong (1987) と Gelfand and Smith (1990) で拡張された。次の標本は、一連の完全条件付き分布からのランダム・ベクトルの部分要素をサンプリングすることによって得られる。その他の MCMC 法には、ハイブリッド MH と棄却サンプリング (Tierney, 1994) と EM アルゴリズムの確率バージョン (Celeux and Diebolt, 1985) などの方法が含まれる。

MCMC を用いた方法は統計学上での応用分析できわめて役に立ち、多くは複雑な、高次元問題に適用するのが難しい伝統的な独立サンプリング法よりもずっと役に立つことが証明されている。MCMC 法は目標密度の正規化定数が分からない場合にも適用することができ、このことは目標 (事後) 密度の正規化定数がほとんどわからないベイズ分析では非常に重要である。加えて、扱いにくい尤度関数を持つモデルをシミュレートできるような MCMC スキームを組み立てることができようになる。これは通常、特に Gibbs サンプルングではパラメータ空間を欠測データ、あるいは潜在変数を含めるように拡張する方法であるデータ拡大によって行う。以下で、代表的なモデルについて Chib and Greenberg (1996) に沿って説明する。

7.1 応用例

MCMC シミュレーション法をいろいろな種類の計量経済モデルに適用する方法を、以下で示すことにする。まず最初に、Gibbs サンプラーをデータ拡大のない場合に適用することができ、シミュレーションが標準的な分布だけからなる単純な例から始める。目的はこの方法の考え方を示し、それを他の場合に適用する方法の役に立つことである。例を示す前に、この後使う事前密度に対する仮定を紹介する。ベクトル β は $\mathcal{N}_k(\beta_0, B_0^{-1})$ に、分散 σ^2 は逆ガンマ分布 $\mathcal{IG}(v_0/2, \delta_0/2)$ に、そして精度行列 Ω^{-1} はウィッシュャート分布 $\mathcal{W}(\rho_0, R_0)$ に従う。事前密度のハイパーパラメータはゼロの添え字を付け、既知であると仮定する。

SUR: Seemingly Unrelated Regression Model

最初の例は SUR モデルで、計量経済学では広く用いられているモデルである。誤差項が正規分布に従うとの仮定の下で、観測されたデータ y_{it} は

$$y_{it} = x'_{it} \beta_i + \epsilon_{it} \quad \epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{pt})' \sim \text{iid. } \mathcal{N}_p(0, \Omega), \\ 1 < i \leq p, 1 \leq t \leq n$$

ここで、 β_i は $k_i \times 1$ のベクトルで Ω は正定符号行列である。各期で観測値を縦にスタックすると、モデルを次ぎのようなベクトル型で書き直すことができる。

$$y_t = X_t \beta + \epsilon_t$$

ここで $y_t = (y_{1t}, \dots, y_{pt})'$, $X_t = \text{diag}(x'_{1t}, \dots, x'_{pt})$, $\beta = (\beta_1, \dots, \beta_p) : k \times 1$ として $k = \sum_i k_i$ である。 $p = 1$ の時、単一方程式正規回帰モデルである。データ $Y_n = (y_1, \dots, y_n)$ の標本に対する最尤推定値は反復過程を通じてのみ得られ、これらの推定量の有限標本分布は取り扱いにくい。対照的に、Gibbs サンプリング・アルゴリズムは、このモデルに対する厳密な小標本ベイズ推定を与える (Percy, 1992; Chib and Greenberg, 1995a)。

(β, Ω^{-1}) に関する事前情報が密度 $\pi(\beta)\pi(\Omega^{-1})$ によって表されると、 β と Ω^{-1} (精度行列) は独立と仮定していることになる。すると、(事前密度と尤度関数の積に比例する) パラメータの事後密度は、次式で与えられる。

$$\pi(\beta)\pi(\Omega^{-1}) \times |\Omega^{-1}|^{n/2} \exp\left[-\frac{1}{2} \sum_{t=1}^n (y_t - X_t \beta)' \Omega^{-1} (y_t - X_t \beta)\right]$$

これは、(正規化定数が未知の) 目標密度であり、ここからシミュレートする。さて、もし β と Ω^{-1} をパラメータの2つのブロックとして扱えば、完全条件付き密度 $p(\beta|Y_n, \Omega^{-1})$ と $p(\Omega^{-1}|Y_n, \beta)$ は簡単にシミュレートできる。特に、すでに述べた事前分布のもとでは

$$p(\beta|Y_n, \Omega^{-1}) \sim \mathcal{N}(\hat{\beta}, B_n^{-1}) \quad \text{と} \quad p(\Omega^{-1}|Y_n, \beta) \sim \mathcal{W}_p(\nu_0 + n, R_n)$$

ここで $\hat{\beta} = B_n^{-1}(B_0 \beta_0 + \sum_{t=1}^n X_t' \Omega^{-1} y_t)$, $B_n = (B_0 + \sum_{t=1}^n X_t' \Omega^{-1} X_t)$, そして $R_n = [R_0^{-1} + \sum_{t=1}^n (y_t - X_t \beta)' \Omega^{-1} (y_t - X_t \beta)]^{-1}$ である。これらの2つの分布を Gibbs アルゴリズムでシミュレートすると、 $\beta^{(i)}$ が周辺密度 $\pi(\beta|Y_n)$, $\Omega^{-1(i)} \sim \pi(\Omega^{-1}|Y_n)$ に従って分布するような標本 $\{\beta^{(i)}, \Omega^{-1(i)}\}$ を生成する。そして $p(\beta^{(i)}, \Omega^{-1(i)})$ は目標 (同時) 密度に従って分布する。抽出されたサンプルは、インポートランス・サンプリング関数や尤度関数を評価することなしに得られる。

Tobit と Probit 回帰モデル

SUR モデルでは、Gibbs サンプラーをモデルのパラメータに直接適用する。しかし直接適用できない場合でも扱いやすい完全条件付き分布のセットは tobit モデルや probit モデルのように、パラメータ空間を直接観測できない潜在データで拡張することによって得ることができる。興味あることに、サンプラーで定義されるパラメータ空間が非常に大きい (probit モデルの場合はサンプルサイズよりも大きい) と、シミュレーションにおけるブロックの数はきわめて小さくなる (tobit モデルでは3、二値 probit モデルでは2である)。

Tobin (1958) の打ち切り回帰モデル (censored regression model) では、観測値 y_i は次のように生成される。

$$z_i \sim \mathcal{N}(x_i \beta, \sigma^2) \quad \text{そして} \quad y_i = \max(0, z_i), \quad 1 < i \leq n$$

n 個の独立な観測値のセットが与えられると、 β と σ^2 に対する尤度関数は次のように表すことができる。

$$\prod_{i \in C} [1 - \Phi(x_i \beta / \sigma)] \prod_{i \in \bar{C}} (\sigma^{-2}) \exp\left[-\frac{1}{2\sigma^2} (y_i - x_i \beta)^2\right]$$

ここで C は一組の打ち切り観測値、そして Φ は標準正規確率変数の累積分布関数である。明らかに、この関数は (事前分布を掛けた後) Gibbs サンプリング・アルゴリズムのなかで簡単に扱えない。計量経済学で Gibbs サンプリングを適用した最初の研究の一つである Chib (1993) は、パラメータ空間が打ち切り観測値に対応する潜在データによって拡張される場合には、非常に

簡単になることを示した。

その理由は、ベクトル $z = \{z_i\}$, $i \in C$ が利用できるためである。 y_z は $n \times 1$ ベクトルで、もし i 番目の観測値がセンサーされていないならば i 番目の要素は y_i で、センサーされていれば z_i である。ブロックを β, σ^2, z とし、それぞれの完全条件付き密度を $p(\beta | Y_n, z, \sigma^2)$, $p(\sigma^2 | Y_n, z, \beta)$, $p(z | Y_n, \beta, \sigma^2)$ とした場合の Gibbs サンプリング・アルゴリズムを考えると、これらの分布はすべて扱いやすいことが分かり、Gibbs シミュレーションはすぐに適用可能である。最初の2つの分布は次のようになる。

$$p(\beta | y_z, \sigma^2) \sim \mathcal{N}_k(\hat{\beta}, (B_0 + \sigma^{-2} X'X)^{-1})$$

そして

$$p(\sigma^2 | y_z, \beta) \sim \mathcal{IG}\left(\frac{\nu_0 + n}{2}, \frac{\delta_0 + \delta_n}{2}\right) \quad (7)$$

ここで $X = (x_1, \dots, x_n)$, $\hat{\beta} = (B_0 + \sigma^{-2} X'X)^{-1}(B_0\beta_0 + \sigma^{-2} X'y_z)$, そして $\delta_n = (y_n - X\beta)'(y_n - X\beta)$ 。他方、潜在データの完全条件付き分布は n 個の独立分布の積に簡略化される。 $p(z | Y_n, \beta, \sigma^2) = \prod_{i \in C} p(z_i | y_i = 0, \beta, \sigma^2)$, ここで

$$p(z_i | y_i = 0, \beta, \sigma^2) \sim \mathcal{I}\mathcal{N}_{(-\infty, 0)}(x_i\beta, \sigma^2), \quad i \in C$$

は、区間 $(-\infty, 0]$ でサポートされる切断正規分布である。この場合(たとえば、 β と σ^2 の分布は、潜在データが与えられたときのセンサーされたデータと独立である)に観測される条件付き独立性への単純化は、通常データ拡大を行ったときに生じ、これはなぜデータ拡大が有効な道具であるかを示している。

データ拡大が有効であることは probit モデルでも明白である。 n 個の独立観測値 $Y_n = \{y_i\}$ が与えられており、各 y_i は $\text{Prob}(y_i = 1) = \Phi(x_i\beta)$ であるベルヌーイプロセスに従う。このモデルやこのクラスに入る他の多くのモデルに対して、Albert and Chib (1993b) は、潜在的なガウス型データを Gibbs サンプリング・アルゴリズムにおけるもう一つの未知パラメータとして導入する単純で、有力な方法を開発した。次式

$$z_i = x_i\beta + u_i, \quad u_i \sim \text{iid. } \mathcal{N}(0, 1), \quad \text{そして} \quad y_i = \mathbb{I}[z_i > 0] \quad (8)$$

は probit モデルを表すことを示した。(データ拡大を含む) Gibbs サンプリング・アルゴリズムは、次の条件付き分布を通じて定義される。

$$p(\beta | Y_n, Z_n) \stackrel{d}{=} p(\beta | Z_n), \quad \text{そして} \quad p(Z_n | Y_n, \beta) \stackrel{d}{=} \prod_{i=1}^n p(z_i | y_i, \beta)$$

ここで、 $Z_n = (z_1, \dots, z_n)$ である。

β の完全条件付き分布は、 y_z を Z_n で置き換え $\sigma^2 = 1$ とした(7)式と同じ形をしている。完全条件付き密度 $p(Z_n | Y_n, \beta)$ は、独立な項の積として分解され、 $y_i = 1$ あるいは $y_i = 0$ に依存している。(8)式から、 $z_i \leq 0$ なら $y_i = 0$, $z_i > 0$ なら $y_i = 1$ となる。そこで

$$p(z_i | y_i) = \begin{cases} 0, & \beta \sim \mathcal{I}\mathcal{N}_{(-\infty, 0)}(x_i\beta, 1) \\ 1, & \beta \sim \mathcal{I}\mathcal{N}_{(0, \infty)}(x_i\beta, 1), \quad 1 \leq i \leq n \end{cases}$$

この MCMC アルゴリズムは、自由度が ν の独立な t リンク関数を持つモデルを推定するように簡単に修正することができる (Albert and Chib, 1993a)。 t 分布は、正規分布と混合するガ

ンマ分布 $\Gamma(\nu/2, \nu/2)$ のスケール混合であるという結果から、パラメータ空間をこのガンマ変数で各変数値ごとにさらに拡張することができる。完全条件付き密度は再び扱いにくい（この考え方を線形回帰に用いたものは、Carlin and Polson, 1990; Geweke, 1993b である）。Albert and Chib (1993b) は ν を未知とし、probit モデルの一般的な頑健性を導いている。

AR(p) 誤差の回帰モデル

この節では MH アルゴリズムを用いたシミュレーションの例を示す。ARMA(p, q) 誤差の回帰モデルの詳細な分析は、Chib and Greenberg (1994) と Marriott, Ravishanker, and Gelfand (1995) に見られる。

次のモデルを考える。

$$y_t = x_t' \beta + \epsilon_t, \quad 1 \leq t \leq n \tag{9}$$

ここで、 y_t はスカラーの観測値である。誤差項は定常な AR(p) プロセス

$$\epsilon_t - \phi_1 \epsilon_{t-1} - \dots - \phi_p \epsilon_{t-p} = u_t \quad \text{or} \quad \phi(L) \epsilon_t = u_t \tag{10}$$

によって生成されていると仮定する。ここで、 $u_t \sim iid. \mathcal{N}(0, \sigma^2)$ として $\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$ はラグオペレータ L の多項式である。定常性の仮定は、 $\phi(L)$ の根が単位円の外にあることを意味する。これは $\phi = (\phi_1, \dots, \phi_p)$ を \mathbb{R}^p の部分空間内にあるように制約する。この制約に従うために、 ϕ の事前分布を $\mathcal{N}(\phi | \phi_0, \Phi_0^{-1}) I_{\phi_0}$ とし、定常区域で切断されている正規分布（そして β と σ^2 に対する標準的な事前分布）を仮定する。このモデルに対する尤度関数は次のように表される。

$$f(Y_n | \beta, \phi, \sigma^2) = \mathcal{O}(\phi) \times (\sigma^2)^{-(n-p)/2} \exp \left[-\frac{1}{2\sigma^2} \sum_{t=p+1}^n (y_t^* - x_t^* \beta)^2 \right]$$

ここで、 $t \geq p+1$ に対して $y_t^* = \phi(L)y_t$, $x_t^* = \phi(L)x_t$, そして

$$\Psi(\phi) = (\sigma^2)^{-p/2} |\Sigma_p|^{-1/2} \exp \left[-\frac{1}{2\sigma^2} (Y_p - X_p \beta)' \Sigma_p^{-1} (Y_p - X_p \beta) \right] \tag{11}$$

は最初の p 個の観測値の（定常な）密度である。以前に、 $Y_p = (y_1, \dots, y_p)'$, $X_p = (x_1, \dots, x_p)'$ として $\Sigma_p = \Phi \Sigma_p \Phi' + e_1(p) e_1(p)'$, そして

$$\Phi = \begin{bmatrix} \phi_{-p} & \phi_p \\ I_{p-1} & 0 \end{bmatrix}$$

$e_1(p) = (1, 0, \dots, 0)'$ として $\phi_{-p} = (\phi_1, \dots, \phi_{p-1})'$ である。

事後密度をこのようにシミュレートする。まず3つの事実がある。最初に、Gibbsの方法は β, ϕ, σ^2 をブロックとしてとるとよい。第2に、 β と σ^2 の完全条件付き分布は、標本密度中の2つの指数項を結びつけた後で簡単に得られる。第3に、 ϕ の完全条件付き密度は、MH アルゴリズムでシミュレートすることができる。次に、いくつか詳細に検討する。

$Y_p^* = Q^{-1} Y_p$ と $X_p^* = Q^{-1} X_p$ と定義し、ここで Q は $Q Q' = \Sigma_p$ を満足する。 $y^* = (y_1^*, \dots, y_n^*)$ とし、 X^* に対しても同様である。最後に、 $e = (e_{p+1}, \dots, e_n)'$ とし、 E は $n-p \times p$ 行列で、 t 番目の行は $(e_{t-1}, \dots, e_{t-p})$ で与えられ、ここで $e_t = y_t - x_t' \beta$, $t \geq p+1$ である。完全条件付き分布は、次のようになることを簡単に示すことができる。

$$\begin{aligned}
 p(\beta | Y_n, \phi, \sigma^2) &\sim \mathcal{N}_k(\hat{\beta}, B_n^{-1}) \\
 p(\phi | Y_n, \beta, \sigma^2) &\propto \Psi(\phi) \times \mathcal{N}_p(\hat{\phi}, \hat{\Phi}_n^{-1}) I_{\phi_0} \\
 p(\sigma^2 | Y_n, \beta, \phi) &\sim \mathcal{IG}\left(\frac{\nu_0 + n}{2}, \frac{\delta_0 + \hat{\delta}_p}{2}\right)
 \end{aligned} \tag{12}$$

ここで $\hat{\beta} = B_n^{-1}(B_0\beta_0 + \sigma^{-2}X^*y^*)$, $B_n = (B_0 + \sigma^{-2}X^*X^*)$, $d_p = \|y^* - X^*\beta\|^2$, $\hat{\phi} = \hat{\Phi}_n^{-1}(\Phi_0\phi_0 + \sigma^{-2}E'e)$ として $\hat{\Phi}_n = (\Phi_0 + \sigma^2E'E)$ である。

β と σ^2 の完全条件付き分布は簡単にシミュレートできる。 ϕ をシミュレートするためには、候補生成密度として $\mathcal{N}_p(\hat{\phi}, \hat{\Phi}_n^{-1})I_{\phi_0}$ をもつ MH 独立連鎖を採用する。すると、MH ステップは次のように実行する。 i 番目のイタレーションで、平均 $\hat{\phi}$ で共分散 $\sigma^{2(i)}\hat{\Phi}_n^{-1}$ を持つ正規密度から、候補 $\phi^{(i+1)}$ を抽出する。もしそれが定常性を満足するなら、この点に確率

$$\min\left\{\frac{\Psi(\phi^{(i+1)})}{\Psi(\phi^{(i)})}, 1\right\}$$

で移動する。そして、そうでない場合は $\phi^{(i+1)} = \phi^{(i)}$ とおく。Chib and Greenberg (1994) このアルゴリズムの収束の十分条件を確かめ、いくつかの実証例を示している。

多重時点における構造変化

時系列データをあつかったモデルでの変化点モデルは重要である。計量経済分析では、構造変化問題といわれ従来は構造変化点があるか否か、あつたとすればいつあつたのかに関心が持たれていた。最初は構造変化点があつただ一つの場合について分析され、多重変化点についての分析は最近行われるようになった。古典的な方法としては変化点が既知の場合についての分析は、Chow (1960), Toyoda (1974), Schmidt and Sickles (1977) が、変化点が未知の一般的な場合は BDE (1975) による CUSUM, CUSUMSQ テストがよく知られている。ベイジアン立場では、最近 Chib (1997) がパラメータ・ベクトルの動きを記述するのに hidden Markov chain に基づく新しいモデル化を提案している。多重変化点を扱うために制約付きの遷移確率行列を使用し、状態変数が現在の値にとどまるか次のより高い値にジャンプするかを離散時間離散状態のマルコフ過程によってモデル化した。モデルはマルコフ連鎖モンテカルロシミュレーション法によって、混合モデルに基づいて推定される。

また、単一変化点についての分析は、Carlin, Gelfand, and Smith (1992) が、多重変化点についての分析は、Inclan (1993), Stephens (1994), Yao (1984), Barry and Hartigan (1993), West, Müller and Esobar (1994), 等によって研究されている。

$Y_T = \{y_1, y_2, \dots, y_T\}$ は時系列で、そのパラメータ ξ_t は未知の変化点 $T_m = \{\tau_1, \dots, \tau_m\}$ で変化し、他は一定であると仮定する。

$$\xi_t = \begin{cases} \theta_1 & \text{if } t \leq \tau_1 \\ \theta_2 & \text{if } \tau_1 < t \leq \tau_2 \\ \vdots & \text{if } \vdots \\ \theta_m & \text{if } \tau_{m-1} < t \leq \tau_m \\ \theta_{m+1} & \text{if } \tau_m < t \leq n \end{cases}$$

ここで $\theta_k \in \mathbb{R}^d$ であり、パラメータ・ベクトル $\Theta = (\theta_1, \dots, \theta_{m+1})$ の推定と、未知の変化点 $T_m = (\tau_1, \dots, \tau_m)$ を検出し、変化点の数が異なるモデルと比較することを目的としている。

この多重変化点モデルに関して多くの研究があり、1つの問題は ξ_t のジャンププロセスを決めることである。ベイジアンでは、変化点で生じる別のレジームのパラメータである τ_k と θ_k の

結合事前分布を求めることと同じである。

通常は階層モデルの形で定式化される：最初に、以前の変化点での変化の確率分布をモデル化し、次に現在の変化点と以前のパラメータで、現在のレジームでのパラメータのプロセスをモデル化し、最後にパラメータと変換点によってデータが生成されるとする。パラメータが与えられたときの、データの分布を考える。\$Y_i=(y_1, \dots, y_i)\$を時点 \$i\$ までの経過とし、\$Y^{i,j}=(y_i, y_{i+1}, \dots, y_j)\$を時間 \$i\$ から \$j\$ までの経過とする。すると、\$(\theta, \gamma)\$ の下でのデータの同時密度は

$$f(Y_n | \theta, \gamma) = \prod_{k=1}^{m+1} f(Y^{\tau_{k-1}, \tau_k} | Y_{\tau_{k-1}}, \theta_k, \tau_k)$$

で表される。ここで \$\tau_0=0, \tau_{m+1}=T\$ である。

各期間での離散的な確率変数 \$s_t\$ を導入し、時点 \$t\$ でのシステムの状態を表し、これは整数値 \$\{1, 2, \dots, m+1\}\$ で、そこから特定の観測値 \$y_t\$ がとられるレジームを示す。すなわち \$s_t=k\$ は、観測値 \$y_t\$ が \$f(y_t | Y_{t-1}, \theta_k)\$ からとられることを示す。変数 \$s_t\$ は、制約付の遷移確率行列を持つ離散時間、離散状態マルコフ・プロセスとしてモデル化されるので変化点モデルになる。1期先遷移確率行列は、次のように表される。

$$P = \begin{pmatrix} p_{11} & p_{12} & 0 & \cdots & 0 \\ 0 & p_{22} & p_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \vdots & 0 & p_{mm} & p_{m, m+1} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

このモデルは、ジャンプ確率 \$p_{ii}(i \leq m)\$ がレジームに依存し、状態の遷移は変化点 \$\gamma_m=(\tau_1, \dots, \tau_m)\$ を識別する一般的な変化点モデルとして見ることができる。\$k\$ 回目の変化は、\$s_{\tau_k}=k\$ と \$s_{\tau_{k+1}}=k+1\$ なら \$\tau_k\$ で起こる。もう一つの見方は hidden マルコフモデルとして見るもので、hidden 状態変数 \$s_t\$ の遷移確率に前に述べたような制約を課す。このモデルの見方は MCMC の計算図式の基礎になっている。パラメータの事前密度 \$\pi(\theta, P)\$ を決め、データ \$Y_n\$ があると事後密度 \$\pi(\theta, P | Y_T) \propto \pi(\theta, P) f(Y_n | \theta, P)\$ に関心を集中する。ここで \$\pi\$ は \$(\theta, P)\$ の事前密度と事後密度を表わすこの事後密度は、パラメータ空間を観測されない状態 \$S_n=(s_1, \dots, s_n)\$ で拡大した後、マルコフ連鎖モンテカルロ法によって要約される。言い換えると、モンテカルロ・サンプリングを事後密度 \$\pi(S_n, \theta, P | Y_T)\$ に適用する。

他方、Kozumi and Hasegawa (1998) はベイズの観点から構造変化を調べる新しい方法を提案した。この方法は Dirichlet 事前分布を用いた階層モデルを基にしている。Dirichlet プロセスの特徴は、構造変化を検出するのに役立つ離散性にあり、それを利用して事前分布を形成する。この論文で彼らは多重変化点を認め、その数が未知であるモデルに対するベイズ・アプローチを提案した。Dirichlet プロセスは魅力があるが、計算が難しいため応用は単純な場合に限定されていたが、最近のマルコフ連鎖モンテカルロ (MCMC) 法の発展により、この点は解消された (Gelfand and Smith, 1990; Besag and Green, 1993 参照)。Dirichlet プロセスは従来は密度の推定に利用されていたが (Escobar, 1994; Escobar and West, 1995), Dirichlet プロセスの離散性がベイズ的なモデル化にも有効な道具であることが最近の多くの研究で示されている (Kuo and Mallich, 1997; Mukhopadhyay and Gelfand, 1997; West, Müller and Esobar, 1994)。この離散性の性質を用いて構造変化を検出する。

実証分析では、1957年から1995年の年データを用いて日本の政府支出に構造変化があるかどうかを検証し、それを CUSUM と CUSUMSQ テストの結果と比較している。

次の単純な線形回帰モデルを考える。

$$y_t = \beta_0 + \beta_1 z_t + u_t, \quad u_t \sim \mathcal{N}(0, \sigma^2), \quad t=1, \dots, T$$

ここで、 y_t と x_t はそれぞれ従属変数と独立変数であり、構造変化を考えるのでパラメータ・ベクトルがある変化点で変化することを認め、次のように書き直す。

$$y_t = \beta_{0t} + \beta_{1t} z_t + u_t, \quad u_t \sim \mathcal{N}(0, \sigma_t^2)$$

ベイズ分析ではパラメータ・ベクトル $\theta_t = (\beta_{0t}, \beta_{1t}, \sigma_t^2)$ の分布が必要であり、前述の Chib (1997) は θ_t の分布を hidden マルコフ連鎖でモデル化した。

$\{\theta_t\}_{t=1}^T$ を Dirichlet プロセスからの標本とする。次に、 θ_t を $k (< T)$ 個のグループに分割し、同じグループ内の θ_t は等しく、異なるグループの θ_t は異なり、 k 個の θ_t の値 $\theta^* = (\theta_1^*, \dots, \theta_k^*)$ は G_0 からサンプルされる。何回構造変化が起こったか、いつ構造変化が起こったのかは、異なる θ^* を調べると分かるので、Dirichlet プロセスの離散性は構造変化点を検出するのに適しているといえる。

パラメータベクトル $\theta_t = (\beta_{0t}, \beta_{1t}, \sigma_t^2)$ の分布 G は未知で、 G の事前分布は Dirichlet プロセスにしたがうと仮定する。

$$y_t \sim \mathcal{N}(x_t' \beta_t, \sigma_t^2), \quad \theta_t \sim G, \quad G \sim \mathcal{DP}(a, G_0)$$

ベイズモデルとするため、残りのパラメータに対し事前分布を決める。基本的な分布 G_0 に対しては、正規分布とガンマ分布とする。

$$dG_0 = \mathcal{N}(\mu, \Sigma) \mathcal{IG}(n_0/2, \tau/2)$$

さらに、階層モデルでの μ, Σ, τ について、事前分布を次のようにおく。

$$\begin{aligned} \mu &\sim \mathcal{N}(\mu_0, V_0) \\ \Sigma^{-1} &\sim \mathcal{W}(\nu_0, \Sigma_0) \\ \tau &\sim \text{Ga}(m_0/2, \tau_0/2) \\ a &\sim \text{Ga}(a_0, b_0) \end{aligned}$$

多重変化点を認めるがその数は未知であるモデルで、誤差分散は変化点にしたがって変化する。ここで $x_t = (1, z_t)'$ と $\beta_t = (\beta_{0t}, \beta_{1t})'$ である。このモデルは階層構造をしており、複雑そうに見えるが Gibbs サンプルングを用いて事後密度の推定を行うことができる。

その他のモデル

今までに述べたモデルに加えて、その他のモデルとしては MCMC 法とデータ拡大を用いる Gibbs サンプルングは特に役に立つ分野である。回帰分析の枠組みでは、欠測値をサンプラーに加えてパラメータの分布から標本を生成する。多対プロビットモデルの重要なクラスは、データ拡大を通じた MCMC シミュレーションによって分析することができ、これらは Albert and Chib (1993a), McCulloch and Rossi (1994), Geweke, Keane, and Runkle (1994) で議論されている。McCulloch and Rossi (1994) にはこの分野での Gibbs サンプルングが役立つ拡張した議論が例とともになされている。変量効果のある一般線形モデルへの応用は、Zeger and Karim (1991) が参考になる。もう一つの重要な分野は混合 (mixture) モデルであり、このモデルではサンプルの各観測値は K 種類の母集団の一つから生じたものである。2 種類のモデルが研究されている。最初のもは、母集団が一つの観測値から次のものへ独立にサンプルされる (Diebolt and Robert, 1994)。2 番目のものは、母集団はマルコフプロセスに従ってサン

ルされ、これはマルコフ・スイッチング・モデル (Albert and Chib, 1993b; Chib, 1993b) と呼ばれる。MCMC法の多様性を示している新しい計量分析の応用は、引き続き多数現れている。本稿で扱わなかった重要な分野で、最近多くの研究が集中的になされている金融時系列(株価, 収益率, 外国為替, 先物等の日次, 週次など高周波データ)の変動に関する計量分析がある。これらには Stochastic Volatility Model, Garch Model, Semi-Non Parametric Model, 線形状態空間モデル, 非線型ガウス・フィルターモデル, 等が含まれる。

8. おわりに

このサーベイでは、ベイズ統計学の計量経済分析への応用に関連するいくつかの重要な問題について議論した。本サーベイは2つの部分に分けられる。ベイジアン計量経済学の考え方と実証分析への応用に関する問題と大きな部分を占める計算問題である。第2章では、ベイジアン計量経済学の簡単な解説を行った。第3章では、応用研究を行う場合におけるベイジアン・パラダイムのいくつかの関連する利点と欠点を述べた。ここで、単位根とEBAに特に焦点を当てたベイジアンの実証分析の簡単な要約を行った。第4章以降では、ベイズ型の実証分析を行う場合に重要な計算問題について、最初に基本的な計算問題について第4章で述べ、ベイジアン研究者が実証研究で利用するいくつかの役に立つ道具について述べた。特に、4.1節では通常出合う問題に接近するためのいくつかの基本的な技術を示し、これらの多くの技法は古典的な研究者が用いているTSPやRATSのようなパッケージと全く同じような方法で、パッケージプログラムを用いて簡単に導入できる。パッケージプログラムが利用できない場合でも、これらの手法のプログラムに必要な時間は、GAUSSやFortranのような言語を使った基本的なプログラムの能力を持っている研究者にとっては最小で済む。第5章では、ある種のクラスの問題を分析するのに適している、もう少し手の込んだベイジアンの手法について述べた。第6章では、現在の計量経済分析を用いた実証研究において主要な手法になったマルコフ連鎖モンテカルロ法について、Metropolis-Hastingアルゴリズム、Gibbsサンプリング、そしてデータ拡大・アルゴリズムについて述べた。MCMC法の理論と実際における最近の発展のサーベイを、計量経済分析への応用に重点を置いて、このアルゴリズムをデータ拡大と結合する方法を用いてベイズ推定への体系的方法を示した。この考え方をセンサー化モデル、離散応答モデル、パネルデータ・モデル、自己回帰モデル、そして時変パラメータモデルの枠組みの中で示したが、この考え方は多くの他の計量モデルに適用できる。

MCMC法が好まれる多くの理由は、他の方法で分析するのが今まで難しかったモデルを分析することを可能にするためである。この方法を利用できるのは、厳密に定式化したモデルと事前分布からなるベイズ分析だけではない。すでに示したように、扱いにくい尤度関数を含む多くのモデルはMCMC法でもシミュレートできる。特に予測、モデルと事前分布の摂動、そしてモデル適合性に関係する様々な推定方程式は、シミュレーションの結果を用いて効果的に記述することができる。

最近では多くの分野での実証研究問題にベイズ法を利用した例が増えてきている。それらの応用例、研究動向などは国際ベイズ分析学会 (IBSA)、4年ごとに開催されるベイズ統計学に関する国際パレンシア研究集会、それに米国統計学会ベイズ統計部会の年次大会などで発表される数多くの論文やProceedingsに見られる。

最後に、このサーベイではできるだけ単純な形で、ベイジアンの手法を用いて実証分析を行うおうとする応用経済分析が利用できる手法を示すことを試みた。したがって、マルコフ連鎖モンテカルロ法で問題になる多くの事柄、特に収束診断法を含む多くの問題点について説明する余裕がなかったので省略した。また、ここで説明した方法で重要な多くの証明や他で示されて

いる式の導出の詳細も省略したので、参考文献を参考にしてほしい。この分野の進展は早いので、最新の時系列分析や金融時系列モデル等の分野についての多量の研究がなされているが、この分野についてのサーベイは次回に譲ることにしたい。また関連する参考文献も完全ではない。

A. ベイズコンピュータ・パッケージについて

ベイズ部分のある標準的な計量分析プログラム

いくつかの標準的なコンピュータパッケージは、よく用いられる経済問題に対して役立つ部分がある。たとえば、RATS マニュアルはベイズ型ベクトル自己回帰モデル (BVAR) に対して RATS を用いたり、SHAZAM はパラメータに線形不等号制約に従う回帰モデルの分析に対するベイズモンテカルロ積分の節に、大きな部分をさいている。

ベイズコンピュータ・パッケージ

A.1 初期に開発されたプログラム

BAYES FOUR 汎用のベイズプログラムで、Prof. Adrian F. M. Smith と彼の同僚である Drs. John Naylor, Allan Skene, Ewart Shaw らによって University of Nottingham で開発された。プログラムではパラメータの期待値としての事後モーメントを計算する。1次元と2次元の周辺事後密度と予測密度も数値的に得られ、グラフィック・ディスプレイも *S-Plus* のプロットライブラリーを利用することによって可能である。10次元ぐらいまでの問題に対しては数値積分(求積法)を用いて、それ以上の次元(20次元ぐらいまで)は非反復モンテカルロ法を用いる。Bayes Four は PC と Unix に対しては、Fortran 77 で書かれており、その他は C あるいは C++ が使える。

BRAP (PCバージョンはPCBRAP) シカゴ大学の Arnold Zellner と同僚が書いた Fortran ベースのプログラムである。正規誤差を持つ線形回帰モデルとして書くことができるいかなるモデルの分析が可能である。加えて、ユーザが与えた関数の(正規あるいは *t* インポートانس関数を用いた)モンテカルロ積分ができる。レビューは Koop (1992) で与えられている。このプログラムのコマンドは TSP のコマンドと非常に似ているので、単純な実証的問題の教育や実行に対しては、最も汎用性があるプログラムである。

BRP CORE の何人かの研究者によって書かれた Fortran ベースのプログラムである。線形モデルの処理と様々な積分手法の両方の広範囲の機能を含んでいる。このサーベイで特に関心があるのは、Poly-*t* 密度を取り扱う数値積分手法である。このパッケージは BRAP よりは汎用性があるが、ユーザ・フレンドリーではない。プログラムの諸機能を使うには、Fortran の知識が必要である。さらに、BRP は適度の込み入った問題に対処するのに適している。

MICRO-EBA GAUSS ベースのプログラムで EBA を分析するためのユタ大学の Richard Fowles が作成したプログラムである。

SEARCH Ed Leamer の推奨する多くの方法を行うプログラムである。

SISAM, MIXIN 2本の Fortran ベースのプログラムで、モンテカルロ積分と混合積分を行う。汎用性はないが、このプログラムが採用している手法は、非常に複雑である。Fortran でプログラムされた poor な動きのする事後分布を持つ研究者に、このプログラムは適している。事後分布が非対称であり、従って t インポートانس関数を用いて扱うのは難しい場合には特に有益である。

A.2 特定目的のプログラム

BATS *Bayesian Analysis of Time Series* プログラムは、カルマンフィルター・モデルを含

んだ動的線形モデル (DLM) と非線形モデルの推定を中心とし、APL で書かれている。もともとは University of Warwick で開発され、West, Harrison and Pole (1987) による論文に記述されている。現在は、強化されたバージョンが Pole, West and Harrison (1994) のベイズ予測と時系列の本で得られる。

bicreg, bic.logit and glib これらの3つのプログラムは、ワシントン大学の Prof. Adrian Raftery が作成した比較的短い S-plus 関数で、ベジアンモデル選択と BIC を用いたモデル不確実性の評価を行う。最初のものは線形回帰モデルでこれを行い、2番目はロジスティック回帰で、3番目は一般線形モデルでこれを行う。

BART *Bayesian Approach to Randomized Trial* は、clinical trial monitoring と indifference zone による分析を行うためのベイズ法をプログラム化した S 関数で、ケンブリッジ大学の MRC Biostatistics Unit の Prof. David Spiegelhalter によって開発された。プログラムは正規尤度を標準事前分布を用いて更新できる。

B/D belief 解析を行うために、Wooff (1992) が作成した [B/D] と呼ばれるプログラムは、*Belief adjusted by Data* の略称で、University of Hull and Durham で開発した Pascal で書かれたプログラムである。未知の数量のネットワーク上で確信を部分的な式を与えると、プログラムは内部的な整合性 (コヒーレンス) をチェックし、他の要素とデータが存在するときにそれぞれを調整する。

Expert System 最近のベイズ的な方法の重要な応用分野はエキスパートシステムである。現在のデータセット集が与えられたときに意思決定者が利用できる情報を自動的に更新する、最も単純で自然な方法はベイズの法則を利用することである。

Hugin

BELIEF 1.2, 2.0

Neural networks software ニューラルネットワークは、エキスパートシステムと密接に関連しており、これらのモデルでの MCMC 法を利用したベイズ学習アルゴリズムが Neal (1994) に書かれており、ソフトウェアが無料で入手できる。

A.3 教育用プログラム

First Bayes Nottingham University の Prof. Tony O'Hagan によって作成され、学生が初等的なベイズ統計学を学習し、理解するのに役立つ使いやすいパッケージである。含まれている機能の一部は次の通り

- ・13種類の標準的な分布族とこれらの混合分布のプロットと要約
 - ・一元配置 ANOVA と単純線形回帰モデルの分析と、位置パラメータの任意の線形結合に対する周辺分布
 - ・ユーザが行ったデータあるいは事前分布に対する変更にしたがって、自動的に事後分布と予測分布の要約統計量を更新する。
 - ・現在の事後分布を事前分布に等しくおくことによるシーケンシャル分析
- プログラムは APL で書かれている。

Minitab Macros Albert (1996) に書かれている Minitab 用のマクロ集。

A.4 マルコフ連鎖モンテカルロプログラム

BUGS は、英国の MRC の Medical Research Unit で開発されたシステムで、*Bayesian inference Using Gibbs Sampling* の略称で、Gibbs サンプリングを手軽に行うために便利なソフトウェアである。D. Spiegelhalter を中心としたグループで開発されている。所与のモデルに対する様々な事前分布：ベルヌーイ分布、ベータ分布、正規分布、2項分布、ディリクレ分布、指数分布、ガンマ分布、対数正規分布、多項分布、正規分布、パ

レート分布, ポワソン分布, 一様分布, ワイブル分布にしたがう乱数を発生させることができ, 完全条件付き分布からのサンプリングを与えることによって多くのベイズ問題を解くことができる。モデルの確率構造(観測値とパラメータ)に対するモデルのスペシフィケーションと確率分布を記述する1組の関数からなる。モデルの記述はモデルが複雑でも簡単である。BUGSを使って既に解析されたモデルと, マニュアルに記述されているのはランダム効果を持つ一般線形モデル, 生存データの線形回帰分析, 空間的な従属性を持つデータとノン・パラメトリック平滑化モデルである。

CODA *Convergence Diagnosis and Output Analysis*: 一般的なMCMCアウトプットやBUGSアウトプットの収束診断を調べるメニュー方式のS-plus関数で, Gelman and Rubin (1992a), Geweke (1992), Ruftery and Lewis (1992) そして Heidelberg and Welch (1983) の要約統計量と収束診断を含んでいる。

Bpois Doss and Narasimhan (1994) によって書かれたベイズポワソン回帰に対する Gibbs サンプリング・プログラム。

MCSim マルコフ連鎖モンテカルロによってベイズ推定を行えるシミュレーションパッケージ。

glimm glimm S関数

BINGO *Bayesian Inference - Numerical, Graphical and Other stuff*

参 考 文 献

- [1] Albert, J and Chib, S. (1993a), "Bayesian analysis of binary and polychotomous response data," *Journal of the American Statistical Association*, 88, 669-679.
- [2] Albert, J and Chib, S. (1993b), "Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts," *Journal of Business and Economics Statistics*, 11, 1-15.
- [3] Bauwens, Luc and M. Lubrano, eds (1995), "Bayesian and classical econometric modelling of time series," *Journal of Econometrics*.
- [4] Bauwens, Luc, W. Polasek and H. van Dijk, eds (1996) "Bayes, Bernoulli, and Basel," *Journal of Econometrics*.
- [5] Berger, J. (1985), *Statistical decision theory and Bayesian analysis*, (second edition), Springer-Verlag: New York.
- [6] Besag, J. and Green, P. J. (1993), "Spatial statistics and Bayesian computation," *Journal of the Royal Statistical Society*, B, 55, 25-37.
- [7] Brown, R. L., Durbin, J, and Evans, J. M. (1975), "Techniques for testing the constancy of regression relationships over times," *Journal of the Royal Statistical Society*, B, 37, 149-192.
- [8] van den Broeck, J., Koop, G., Osiewalsky, J. and Steel, M. (1994), "Stochastic frontier models: a Bayesian perspective," *Journal of Econometrics*, 61, 273-303.
- [9] Carlin, B. P., A. E. Gelfand, and A. F. M. Smith (1992), "Hierarchical Bayesian Analysis of change point problems," *Journal of the Royal Statistical Society C*, 41, 389-405.
- [10] Carlin, B. P. and Louis, T. A., (1996) *Bayes and empirical Bayes methods for data analysis*, Chapman & Hall.
- [11] Carlin, B. P. and N. G. Polson (1991), "Inference for nonconjugate Bayesian models using the Gibbs sampler," *Canadian Journal of Statistics*, 19, 399-405.
- [12] Carlin, B. P., N. G. Polson and D. S. Stoffer (1992), "A Monte carlo Approach to nonnormal and nonlinear state-space modeling," *Journal of American Statistical Association*, 87, 493-500.
- [13] Carter, C. and R. Kohn, (1994), "On Gibbs sampling for state space models," *Biometrika*, 81, 541-554.
- [14] Casella, G., and George, E. (1992), "Explaining the Gibbs sampler," *American Statistician*, 46, 167-174.
- [15] Celeux, G. and Diebolt, J. (1985), "The SEM algorithm: A probabilistic teacher algorithm derived from the EM algorithm for the Mixture problem," *Computational Statistics Quarterly*, 2, 73-82.
- [16] Chib, S. (1992), "Bayes Inference for the Tobit censored regression models," *Journal of Econometrics*,

- 51, 79-99.
- [17] Chib, S. (1993), "Bayes regression with autoregressive errors: a Gibbs sampling approach," *Journal of Econometrics*, **58**, 275-294.
- [18] Chib, S. (1995), "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association*, **90**, 1313-1321.
- [19] Chib, S. (1996), "Calculating posterior distributions and modal estimates in Markov mixture models," *Journal of Econometrics*, **75**, 79-98.
- [20] Chib, S. (1998), "Estimation and comparison of multiple change-point models," *Journal of Econometrics*, **86**, 221-241.
- [21] Chib, S. and Greenberg, E. (1994), "Bayes inference via Gibbs sampling in regression models with ARMA (p, q) Errors," *Journal of Econometrics*, **64**, 183-206.
- [22] Chib, S. and Greenberg, E. (1995a), "Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models," *Journal of Econometrics*, **68**, 339-360.
- [23] Chib, S. and Greenberg, E. (1995b), "Understanding the Metropolis-Hastings algorithm," *American Statistician*, **49**, 327-335.
- [24] Chib, S. and Greenberg, E. (1996), "Markov chain Monte Carlo simulation methods in econometrics," *Econometric Theory*, **12**, 409-431.
- [25] Chow, G. C. (1960), "Tests of equality between sets of coefficients in two linear regressions," *Econometrica*, **28**, 591-605.
- [26] Cooley, T. and LeRoy, S. (1986), "What will take the con out of econometrics: a reply to McAleer, Pagan and Volker," *American Economic Review*, **76**, 504-507.
- [27] Cooley, T. and LeRoy, S. (1991), "Identification and estimation of money demands," *American Economic Review*, **71**, 825-844.
- [28] DeJong, D. (1992), "Co-integration and trend-stationarity in macroeconomic time series: Evidence from the likelihood function," *Journal of Econometrics*, **52**, 347-370.
- [29] DeJong, D. and N. Shephard (1992), "The Simulation smoother for time series Models," *Biometrika*, **82**, 339-350.
- [30] DeJong, D. N. and C. H. Whiteman (1989a) "Trends and random walks in macroeconomic time series: A reconsideration based on the likelihood principle," *Journal of Monetary Economics*.
- [31] DeJong, D. and Whiteman, C. (1989b), "Trend and cycles as unobserved components in GNP: A Bayesian perspective," *Proceedings of the American Statistical Association, Business and Economics Section*, 63-70.
- [32] DeJong, D. and Whiteman, C. (1991a), "The temporal stability of dividends and stock prices: Evidence from the likelihood function," *American Economic Review*, **81**, 600-617.
- [33] DeJong, D. and Whiteman, C. (1991b), "Reconsidering trends and random walks in macro-economic time series," *Journal of Monetary Economics*, **28**, 221-254.
- [34] DeJong, D. and Whiteman, C. (1991c), "The Case for trend-stationarity is stronger than we thought," *Journal of Applied Econometrics*, **6**, 413-421.
- [35] DeJong, D. and Whiteman, C. (1996), "Trend and cycles as unobserved components in GNP: A Bayesian Perspective," in *Advances in Econometrics*, eds by T. B. Fomby and R. C. Hill Vol. 11 B, 29-49.
- [36] DeJong, D., Nankervis, J. Savin, N. and Whiteman, C. (1992), "The power problems of unit root tests in time series with autoregressive Errors," *Journal of Econometrics*, **53**, 323-343.
- [37] DeJong, D. and Whiteman, C. (1993), "Estimating moving average parameters: Classical pileups and Bayesian posteriors," *Journal of Business and Economic Statistics*, **11**, 311-317.
- [38] Dempster, Laird and Rubin (1977), "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society B*, **39**, 1-38.
- [39] Dickey, D. A. and Fuller, W. A. (1979), "Distribution of the estimators for autoregressive time-series with a unit root," *Journal of the American Statistical Association*, **74**, 427-431.
- [40] Doan, T. (1988), *RATS: Users Manual, version 4.*, Estima, Evanston, Illinois.
- [41] Dorfman, J.H. (1997), *Bayesian Economic through Numerical Method*, Springer Verlag.
- [42] Dreze, J. (1977), "Bayesian regression Analysis using Poly- t densities," *Journal of Econometrics*, **6**, 329-354.

- [43] Escobar, M. D. (1994) "Estimating normal means with a Dirichlet process prior," *Journal of the American Statistical Association*, **89**, 268-277.
- [44] Escobar, M. D. and West, M. (1995) "Bayesian density estimation and inference using mixture," *Journal of the American Statistical Association*, **89**, 577-587.
- [45] Frühwirth-Schnatter, S. (1994) "Data augmentation and dynamic linear models," *Journal of Time Series Analysis*, **15**, 183-202.
- [46] Gamerman, D. (1997), *Markov chain Monte Carlo : Stochastic simulation for Bayesian inference*, Chapman & Hall.
- [47] Gelfand, A., Hills, S. E., Racine-Poon, A., and Smith, A. F. M., (1990), "Illustration of Bayesian inference in normal data models using Gibbs sampling," *Journal of the American Statistical Association*, **85**, 972-985.
- [48] Gelfand, A. and Smith, A. F. M. (1990), "Sampling-based approaches to calculating Marginal densities," *Journal of the American Statistical Association*, **85**, 398-409.
- [49] Gelfand, A., Smith, A. F. M., and T. M. Lee (1992), "Bayesian analysis of constrained parameter and truncated data problems," *Journal of the American Statistical Association*, **85**, 398-409.
- [50] Gelman, A. and Rubin, D. (1992), "A Single series from the Gibbs sampler provide a false sense of security," *Bayesian Statistics 4*, J. M. Bernardo, J. M. Berger, A. P. Dawid, and A. F. M. Smith (eds), Oxford University Press, Oxford.
- [51] Geman, S. and Geman, D. (1984), "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721-741.
- [52] Geweke, J., (1986), "Exact inference in the inequality constrained normal linear regression models," *Journal of Applied Econometrics*, **1**, 127-141.
- [53] Geweke, J., (1988), "Antithetic accelerations of Monte carlo integration in Bayesian econometrics," *Journal of Econometrics*, **38**, 73-90.
- [54] Geweke, J., (1989a), "Bayesian inference in econometric models using Monte carlo integration," *Econometrica*, **57**, 1317-1340.
- [55] Geweke, J., (1989b), "Modeling with normal polynomial expansions," in *Economic Complexity : Chaos, Sunspots, Bubbles and Nonlinearity*, edited by W. Barnett, J. Geweke, and K. Shell, Cambridge University Press.
- [56] Geweke, J. (1992a), "Priors for macroeconomic time series and their applications," *Econometric Theory*, **10**, 609-632.
- [57] Geweke, J., (1992b), "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments," *Bayesian Statistics 4*, J. M. Bernardo, J. M. Berger, A. P. Dawid, and A. F. M. Smith (eds), Oxford University Press, Oxford, 169-193.
- [58] Geweke, J. (1997) "Posterior Simulators in Econometrics," D. M. Kreps and K. F. Wallis eds., *Advances in economics and Econometrics : Theory and Applications*, Seventh World Congress Vol. 3, Cambridge University Press.
- [59] Gilks, W. R., Richardson, S., and Spiegelhalter, D. J., (1996), "Introducing Markov chain Monte Carlo," in Gilks, Richardson, and Spiegelhalter eds. *Markov Chain Monte Carlo in Practice*, Chapman & Hall.
- [60] Gilks, W. R. and Wild, P. (1992), "Adaptive rejection sampling for Gibbs sampling," *Journal of the Royal Statistical Society, C* **41**, 337-348.
- [61] Hastings, W. K. (1970), "Monte carlo sampling methods using Markov chains and their applications," *Biometrika*, **57**, 97-109.
- [62] Hatanaka, M. (1996), *Time-series-based econometrics : Unit roots and cointegration*, Advanced Texts in Econometrics, Oxford University Press.
- [63] Heidelberger, P. and Welch, P. D. (1983), "Simulation run length control in the presence of an initial transient," *Operations Research*, **31**, 1109-44.
- [64] Hop, J. and van Dijk, H. (1990), "SISAM and MIXIN : Two algorithms for the computation of posterior moments and densities using Monte Carlo integration," Econometric Institute, Erasmus University Report 9031/A.
- [65] Inflan, C. (1993), "Detection of multiple changes of variance using posterior odds," *Journal of Business and Economic Statistics*, **11**, 289-300.
- [66] Jacquier, E., Polson, N. and Rossi, P. (1994), "Bayesian analysis of stochastic volatility models (with

- discussion)," *Journal of Business and Economic Statistics*, **12**, 371-417.
- [67] Jeffreys, H. (1961), *Theory of Probability*, 3rd ed., Oxford University Press, London.
- [68] Kadane, J. (1984), *Robustness of Bayesian Analysis*, North-Holland, Amsterdam.
- [69] Kadane, J., Dickey, J. Winkler, R., Smith, W. and Peters, S. (1980), "Interactive elicitation of option for a Normal linear models," *Journal of the American Statistical Association*, **75**, 845-854.
- [70] Kass, R. E., L. Tierney, and J. B. Kadane (1988), "Asymptotics in Bayesian computation (with discussion)," In *Bayesian Statistics 3*, J. M. Bernardo, M. H. DeGroot, A. P. Dawid, and A. F. M. Smith, eds. Oxford University Press, 261-278.
- [71] Kass, R. E., L. Tierney, and J. B. Kadane (1990), "The Validity of posterior expansions based on Laplace's method," in S. Geisser, J. S. Hodges, S. J. Press, and A. Zellner eds. *Bayesian and Likelihood Methods in Statistics and Econometrics*, 473-487.
- [72] Kass, R. E., Carlin, B. P., gelman, A., and Neal, R. M. (1998), "Markov Chain Monte Carlo in Practice: A Roundtable Discussion," *American Statistician*, **52**, 93-100.
- [73] Kloeck, T. and van Dijk, H. (1978), "Bayesian estimates of equation system Parameters: an Application of integration by Monte carlo," *Econometrica*, **46**, 1-20.
- [74] Koop, G. (1992a), "Objective' Bayesian unit root tests," *Journal of Applied Econometrics*, **7**, 65-82.
- [75] Koop, G., (1992b), "Aggregate shocks and macroeconomic fluctuations: a predictive approach," *Journal of Applied Econometrics*, **7**, 395-412.
- [76] Koop, G., (1992c), "Intertemporal properties of real output: a Bayesian approach," *Journal of Business and Economic Statistics*, **9**, 253-265.
- [77] Koop, G. and Poirier, D. (1993), "A Bayesian analysis of Logit models using natural conjugate priors," *Journal of Econometrics*, **56**, 323-340.
- [78] Koop, G. and Steel, M. (1994), "A Decision theoretic analysis of the unit root hypothesis using mixtures of elliptical models," *Journal of Business and Economic Statistics*, **12**, 95-107.
- [79] Kozumi, Hideo and Hasegawa, H. (1998), "A Bayesian analysis of structural changes with an application to displacement effect," paper presented at Japan Statistical Association annual meeting.
- [80] Kuo, L. and Mallick, B. K. (1997), "Bayesian semiparametric inference for the accelerated failure time model," *Canadian Journal of Statistics*,
- [81] Leamer, E., (1985), "Sensitivity analysis would help," *American Economic Review*, **75**, 308-313.
- [82] Leamer, E., (1983), "Let's take the con out of Econometrics," *American Economic Review*, **73**, 31-43.
- [83] Leamer, E., (1978), *Specification Searches: Ad-hoc inference with non-experimental data*, Wiley, New York.
- [84] Lee, Peter (1997), *Bayesian statistics: An introduction*, 2nd. Ed. Arnold.
- [85] Marriott, J., N. Ravishanker, and A. E. Gelfand (1995), "Bayesian analysis of ARMA processes: Complete sampling-based inferences under full likelihood," in D. Berry, K. Chaloner, and J. Geweke eds. *Bayesian Statistics and Econometrics: Essays in Honor of Arnold Zellner*.
- [86] Marsh, T. A. and R. C. Merton (1986) "Dividend variability and variance bounds tests for the rationality of stock market prices," *American Economic Review*, **76**, 483-494.
- [87] McAleer, M., Pagan, A., and Volker, P. (1985), "What will take the con out of Econometrics?," *American Economic Review*, **75**, 293-307.
- [88] McCulloch, R. and Rossi, P., (1992), "Bayes factors for nonlinear hypothesis and likelihood distributions," *Biometrika*, **79**, 663-676.
- [89] McCulloch, R. and Rossi, P. (1994), "An Exact likelihood analysis of the multinomial probit models," *Journal of Econometrics*, **64**, 207-240.
- [90] McCulloch, R. and Tsay, R., (1994a), "Statistical analysis of economic time series via Markov switching models," *Journal of Time Series Analysis*, **15**, 235-250.
- [91] McCulloch, R. and Tsay, R., (1994b), "Bayesian inference of trend- and difference- stationarity," *Econometric Theory*, **10**, 596-608.
- [92] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E (1953), "Equations of state calculations by fast computing machines," *Journal of Chemical Physics*, **21**, 1087-1092.
- [93] Mukhopadhyay, S. and Gelfand, A. E. (1997), "Dirichlet process mixed generalized linear model," *Journal of the American Statistical Association*, **92**, 633-639.
- [94] Müller, P. (1991), "A Generic approach to posterior integration and Gibbs sampling," Technical Report

- 91-09, Department of Statistics, Purdue University.
- [95] Naylor, J. and Shaw, J. (1991), *Bayes Four User Guide*, Nottingham Polytechnic, manuscript.
- [96] Naylor, J. and Smith, A., (1982), "Applications of method for the efficient consumption of posterior distributions," *Applied Statistics*, **31**, 214-225.
- [97] Naylor, J. and Smith, A. (1988), "Econometric illustrations of novel numerical integration strategies for Bayesian inference," *Journal of Econometrics*, **38**, 103-126.
- [98] Nelson, C. and Plosser, C. (1982), "Trends and random walks in macroeconomics time series," *Journal of Monetary Economics*, **10**, 139-162.
- [99] Phillips, P. (1991), "To criticize the critics: an objective Bayesian analysis of stochastic trend (with discussion)," *Journal of Applied Econometrics*, **6**, 333-364.
- [100] Phillips, P. and Ploberger, W. (1991), "Time series modelling, with a Bayesian frame of reference: I. Concepts and Illustrations, Cowles Foundation Discussion Paper No. 980.
- [101] Phillips, P. and Perron, P. (1988), "Testing for unit roots in time series regression," *Biometrika*, **71**, 599-607.
- [102] Poirier, D. (1988), "Frequentist and subjectivist perspectives on the problems of model building in Econometrics, (with Discussion)", *Journal of Economic Perspectives*, **2**, 121-170.
- [103] Poirier, D. ed. (1991), "Bayesian empirical studies in economics and finance," *Journal of Econometrics*,
- [104] Poirier, D. (1995), *Intermediate Statistics and Econometrics: A Comparative Approach*, MIT Press.
- [105] Raftery, A. and Lewis, S. (1992), "How many iterations in the Gibbs sampler?" *Bayesian Statistics 4*, J. M. Bernardo, J. M. Berger, A. P. Dawid, and A. F. M. Smith (eds), Oxford University Press, Oxford, 115-130.
- [106] Richard, J. and Tompa, H. (1980), "On the evaluation of poly-*t* density functions," *Journal of Econometrics*, **12**, 335-352.
- [107] Ripley, B. (1987), *Stochastic Simulation*, New York, Wiley.
- [108] Rubin, D.B., (1987), 'Comment on "The calculation of posterior distributions by data augmentation" by M. A. Tanner and W. H. Wong, *Journal of the American Statistical Association*, **82**, 543-546.
- [109] Schmidt, P. and P. C. B. Phillips (1989) "Testing for a unit root in the presence of deterministic trends," Cowles Foundation Discussion Paper No. 933, New Haven, Ct.
- [110] Schmidt, P. and P. C. B. Phillips (1992) "LM tests for a unit root in the presence of deterministic trends," *Oxford Bulletin of Economics and Statistics*, **54**, 257-87.
- [111] Schmidt, P. and Sickles, R. (1977), "Some further evidence on the use of the Chow test under heteroscedasticity," *Econometrica*, **45**, 1293-1298.
- [112] Schotman, P. and van Dijk, H. (1991a), "A Bayesian analysis of the unit root in real exchange rates," *Journal of Econometrics*, **49**, 195-238.
- [113] Schotman, P. and van Dijk, H., (1991b), "On Bayesian routes to unit roots," *Journal of Applied Econometrics*, **6**, 387-402.
- [114] Schwert, G. (1989), "Tests for unit roots: a Monte carlo Investigation," *Journal of Business and Economic Statistics*, **7**, 147-160.
- [115] Shiller (1981) "Do stock prices move too much to be justified by subsequent changes in dividends?" *American Economic Review*, **71**, 421-436.
- [116] Sims, C. (1988), "Bayesian skepticism on unit root econometrics," *Journal of Economic Dynamics and Control*, **12**, 463-474.
- [117] Sims, C. and Uhlig, H. (1991), "Understanding unit rooters: a helicopter tour," *Econometrica*, **59**, 1591-1599.
- [118] Smith, A. F. M. and Gelfand, A. E. (1992), "Bayesian statistics without tears," *American Statistician*, **46**, 84-88.
- [119] Smith, A., Skene, A., Shaw, J., Naylor, J. and Dransfield, M. (1985), "The implementation of the Bayesian paradigm," *Communications in Statistics*, **14**, 1079-1102.
- [120] Stephens, D. A. (1994), "Bayesian retrospective multiple-change point identification," *Applied Statistics*, **43**, 159-178.
- [121] Tanner, M. (1991), *Tools for Statistical Inference*, Springer-Verlag, Lecture Notes in Statistics Series, edited by J. Berger, S. Fienberg, J. Gani, K. Krickberg, I. Olkin and B. Singer.
- [122] Tanner, M. and Wong, W. (1987), "The Calculation of posterior distributions by data augmentations,

- (with Discussion)," *Journal of the American Statistical Association*, **82**, 528-550.
- [123] Tierney, L. (1991), "Exploring posterior distributions using Markov chains," in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, E. M. Keramides and S. M. Kaufman (eds.), Interface Foundation of North America.
- [124] Tierney, L. (1994), "Markov chains for exploring posterior distributions," *Annals of Statistics*, **22**, 1701-1728, with discussion.
- [125] Tierney, L. and Kadane, J. (1986), "Accurate approximations for posterior moments and marginal distributions," *Journal of the American Statistical Association*, **81**, 82-86.
- [126] Tierney, L., Kass, R. E., and Kadane, J. (1987), "Interactive Bayesian analysis using accurate asymptotic approximations," In *Computing Science and Statistics: 19th Symposium on Interface*, R. M. Heiderberger, 15-21.
- [127] Tierney, L., Kass, R. E., and Kadane, J. (1989), "Fully exponential Laplace approximations to expectations and variances of nonpositive functions," *Journal of the American Statistical Association*, **84**, 710-716.
- [128] Thisted, R. A. (1988), *Elements of Statistical Computing*, New York, Chapman & Hall.
- [129] Toyoda, T. (1974), "The use of the Chow test under heteroscedasticity," *Econometrica*, **42**, 601-608.
- [130] Tsurumi, H. and Wago, H. (1996), "A Bayesian analysis of unit root and cointegration with an application to Yen-Dollar exchange rate model," in *Advances in Econometrics*, eds by T. B. Fomby and R. C. Hill, Vol. 11 B, 51-86, JAI Press.
- [131] Yao, G. C. G. (1984), "Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches," *Annals of Statistics*, **12**, 1434-1447.
- [132] West, M., Müller, P., and Escobar, M. D. (1994), "Hierarchical priors and mixture models, with application in regression and density estimation," In *Aspects of Uncertainty: A Tribute to D. V. Lindley*, ed. by Smith, A. F. M. and Freeman, P., 363-386, New York: Wiley.
- [133] White, K., Haun, S. and Horsman, N. (1987), "SHAZAM: Users Reference Manual, Version 7."
- [134] Zangari, P. J. and H. Tsurumi (1996), "A Bayesian analysis of censored autocorrelated data on exports of Japanese passenger cars to United States," in *Advances in Econometrics*, eds. by T. B. Fomby and R. C. Hill, Vol. 11 A, 111-143, JAI Press.
- [135] Zeger, S. and Karim, M. (1991), "Generalized linear models with random Effects: a Gibbs sampling approach," *Journal of the American Statistical Association*, **86**, 79-86.
- [136] Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons, Inc., New York.
- [137] Zellner, A., (1983), "Applications of Bayesian analysis in econometrics," *The Statistician*, **32**, 23-34.
- [138] Zellner, A., (1985), "Bayesian Econometrics," *Econometrica*, **53**, 253-269.
- [139] Zellner, A., (1986), "On assessing prior distributions and Bayesian regression analysis with g -prior distributions," in *Bayesian Inference and Decision techniques*, edited by P. Goel and A. Zellner, Elsevier Science, London.
- [140] Zellner, A. and Rossi, P. (1984), "Bayesian analysis of dichotomous quantal response models," *Journal of Econometrics*, **25**, 365-393.
- [141] Zivot, E., (1992) "Bayesian analysis of the unit root hypothesis within an unobserved component model," *mimeograph*.
- [142] Zivot, E. and P. C. B. Phillips, (1991) "A Bayesian analysis of trend determination in economic time series," Cowles Foundation Discussion Paper, No. 1002, New Haven, CT.
- [143] 伊庭幸人 (1996), 「マルコフ連鎖モンテカルロ法とその統計学への応用」統計数理, **44**, 49-84.
- [144] 大森裕浩 (1996), 「マルコフ連鎖モンテカルロ法」千葉大学経済研究, **10**, 237-287.
- [145] 繁辨算男 (1995), 「意志決定の認知統計学」朝倉書店, 付録 137-148.
- [146] 下平英寿 (1997), 「ベイズ的方法におけるMCMCの利用」応用統計学会チュートリアル講演予稿集, 1-10.
- [147] 伴 金美 (1991), 「マクロ計量モデル分析」有斐閣.