

論文審査の要旨及び担当者

No. 1

| 報告番号 | 甲 第 号 | 氏 名 | 中北 誠 |
|--|-------|---------------------------------|------|
| 論文審査担当者 | 主 査 | ：星野 崇宏 (慶應義塾大学経済学部教授 博士 (経済学)) | |
| | 副 査 | ：長倉 大輔 (慶應義塾大学経済学部教授 Ph.D.) | |
| | | ：中妻 照雄 (慶應義塾大学経済学部教授 Ph.D.) | |
| | 面接担当 | ：沖本 竜義 (慶應義塾大学経済学部教授 Ph.D.) | |
| | | ：片山 翔太 (慶應義塾大学経済学部准教授 博士 (理学)) | |
| (論文審査の要旨) | | | |
| 論文題名：ESSAYS ON BAYESIAN COMPUTATION WITH ANCILLARITY-SUFFICIENCY INTERWEAVING STRATEGY | | | |
| 1. 論文の概要 | | | |
| <p>ベイズ統計学において事後分布とそれに関係する各種統計量を求めるためには高次元の多重積分を評価しなければならないが、解析的に求められる限られた場合を除き、この計算負荷は極めて重いため、長らくベイズ統計学は実証分析で広く応用されることはなかった。しかし、1980年代後半に始まった Markov Chain Monte Carlo (MCMC) 法の普及により、この問題は原理的には解消され、ベイズ統計学の応用分野は飛躍的に拡大した。</p> <p>しかし、「原理的に計算できる」ということは必ずしも「実用的な時間内に計算が完了する」ことを意味しない。MCMC 法の応用においては、事後分布に収束したとみなされる乱数系列を多重積分の評価に利用するが、この乱数系列が事後分布の台を満遍なく巡回しない限り、パラメータの点推定や区間推定の妥当性が保証されない。さらに乱数系列の自己相関関数の減衰が極めて緩やかである場合、多重積分の近似精度が落ちることも知られている。したがって、実用に耐えうる効率的な MCMC 法には、事後分布の台の中を自由自在に動き回る乱数系列を生成できることが要求される。</p> <p>以上の意味で効率的な乱数系列を生成するための手法として、Yu and Meng (2011) が提案した Ancillarity-Sufficiency Interweaving Strategy (ASIS) が知られている。ASIS は潜在変数を含むモデルに適用される MCMC 法の一つである Data Augmentation を改良する手法である。Data Augmentation の定式化には、モデル内で潜在変数を規定する部分にパラメータを含む Sufficient Augmentation (SA) と潜在変数を規定する部分にパラメータを含まない Ancillary Augmentation (AA) がある。SA と AA のどちらが効率的であるかは推定すべき未知のパラメータの値に依存するため、事前にどちらを使うべきかを定めることはできない。そこで、Yu and Meng (2011) は SA と AA を交互に適用することで Data Augmentation の効率向上を図る ASIS を提案した。</p> <p>本論文の主たる目的は、計量経済分析で広く使われるパネルデータ回帰モデルと非線形非正規状態空間モデルの一種である Stochastic Volatility (SV) モデルに対する ASIS の有効性を確認するとともに、各モデルの実データへの応用を行うことである。本論文は 5 つの章からなるが、第 1 章は論文の導入と要約、第 5 章は全体のまとめと今後の展望について述べているだけなので、第 2 章から第 4 章の内容について簡単に紹介する。</p> | | | |

第2章では、Yu and Meng (2011) による ASIS の原理について説明した後、パネルデータ回帰モデルに対する ASIS のアルゴリズムを導出し、Vella and Verbeek (1998) が使用した賃金パネルデータに適用して ASIS の有効性を確認している。本章で扱うパネルデータ回帰モデルはクロスセクションの個体効果を含む典型的なものであり、変量効果モデルあるいは共通の事前分布を仮定した固定効果モデルと解釈される。このモデルに対しては MCMC 法の一つである Gibbs Sampler を容易に実行できることが知られているが、本章の実証分析でも示されるように、従来手法では特に個体効果に関するパラメータの乱数系列で自己相関関数が 1 から減衰しにくいという問題を抱えている。しかし、ASIS を適用することで乱数系列の自己相関関数は急激に減衰するようになり、571 個に及ぶ全てのパラメータの乱数系列が収束判定を通ることが確認された。

第3章では、第2章で導出したパネルデータ回帰モデルのための ASIS を 2 種類の個体効果を持つアンバランス・パネルデータのモデルに拡張し、日本中央競馬会 (JRA) の競馬レースに関するデータを使用して競走馬の能力と騎手の技能の評価を試みている。このパネルデータ回帰モデルでは、各レースにおける競走馬の平均速度を被説明変数とし、4000 頭以上の競走馬と 140 名以上の騎手に対する個体効果、性別、年齢、体重などの競走馬の特性、馬場の状態などを説明変数としている。多数のパラメータを含むモデルにも関わらず、ASIS の効果により全ての乱数系列で収束が確認された。モデルの推定結果より、レースにおける競走馬の平均速度は、競走馬の特性や馬場の状態だけではなく、個体効果として推定された競走馬の能力と騎手の技能からも影響を受けることがわかった。特に競馬界で優秀であると認知されている競走馬や騎手に対して平均速度に対する個体効果が大きくなる傾向がみられた。

第4章では、資産収益率の SV モデルをベイズ推定するための新しいアルゴリズムを提案し、これを株価指数 (TOPIX) の高頻度金融時系列データに応用している。この新アルゴリズムでは、第2章と第3章で有効性を確認した ASIS を活用することで乱数系列の収束の促進を図っている。第4章が提案する SV モデルの特筆すべき点としては、(i) 資産収益率の分布として歪みと厚い裾を持つ一般化双曲分布を仮定していること、(ii) 取引時間内に観測されるボラティリティの周期性 (日中季節性) を Bernstein 多項式で他のパラメータと同時に推定していること、(iii) 資産収益率とボラティリティの間の負の相関 (レバレッジ効果) に対して従来とは異なる定式化を行うことで MCMC 法のアルゴリズムを改良していること、などがあげられる。

2. 論文の評価

次に本論文の主たる研究成果である第2章～第4章に対する評価を簡潔に述べる。

第2章で示されたパネルデータ回帰モデルの MCMC 法の改良は、Gelfand et al. (1995) などの先行研究でも提案されているが、ASIS を適用した事例は知られておらず、後述のように荒削りの部分は残っているものの、新規性の高い研究であるといえる。

第3章で行った競馬データの分析であるが、個々の競走馬の能力の評価を国内競馬について試みた研究としてはおそらく初めてのものであろう。さらに騎手の技能の評価に踏み込んでいる点にも新規性が認められる。参考までに、第3章の元となった論文は、当初は国内の統計学関連学会での報告や学術誌への投稿のために日本語で書かれたもの

であったが、その後修正を加えて英語の論文に変更し、海外のスポーツ統計学に関する学術誌に投稿する予定である。使用している統計分析手法とデータの斬新さから判断すると、将来公刊されることが期待される内容である。

第4章で提案されたSVモデルのベイズ推定のための新アルゴリズムであるが、先行研究である Nakajima and Omori (2009) などと比べると、ASIS の適用や日中季節性の導入に加えて、一般化双曲分布として Skew t 分布だけでなく Variance Gamma 分布にも対応できること、対数ボラティリティの AR(1)過程の分散共分散行列の逆行列が帯行列となることを利用して Block Sampler のアルゴリズムを簡略化していること、従来手法ではモードが1に近いベータ分布を AR(1)係数の事前分布として使うことが通例であるが（そう仮定しないと MCMC 法が破綻する）ASIS を使うことで AR(1)係数にフラットな事前分布（一様分布）を使用できること、などの新規性が見られる。なお第4章の元となった論文は、*Journal of Risk and Financial Management*において” Bayesian analysis of intraday stochastic volatility models of high-frequency stock returns with skew heavy-tailed errors” として 2021年3月に公刊されている。

しかし、口頭試問では審査委員より本論文に対して幾つかの疑問や課題が指摘された。

例えば、ある審査委員から、第2章や第3章でパラメータ数が大変多いモデルを推定しているが、これがうまくできているのはなぜか、という質問があった。中北君によると、「ここで使用しているモデルはベイズ統計学において階層ベイズモデルと呼ばれるものであり、その特性のおかげで推定がうまく機能している。」ということであった。

また、別の審査委員からは、第2章の結果から ASIS によって乱数系列の収束が改善するのはわかったが、それが点推定などの精度の改善につながるのか、そのためには実データだけでなく真のモデルが既知である人工データに対して適用してみる必要があるのではないかと、OLS などとも比較してみてもどうか、という意見が出された。中北君の回答は「乱数系列の収束の速さからみて、同じ数の乱数を生成して点推定を行うという状況下であれば ASIS によって推定精度は向上すると予想されるが、具体的な検証は今後の課題である。」というものであった。

一方、他の審査委員からは、第2章で使用している単純な Gibbs sampler 以外の手法とも比較をした方がよいのではないかと、例えば潜在変数である個体効果を積分して消去したモデルなどに MCMC 法を適用することもできるのではないかと、という意見も出た。中北君によると、これもまた将来の研究に残された課題であるということであった。

さらに第3章で分析した競走馬と騎手の能力評価モデルの応用に関連して、この手法で競馬のレース結果を予想できるか、という質問も出た。中北君はまだ試してはいないということであるが、予想に使うためにはレースの開始前に入手可能な情報だけに依存する形にモデルを修正する必要があるという回答であった。

審査委員から出された疑問や課題は本論文の拡張の方向性を示唆する性格のものであり、本論文の学術的な価値や貢献を損なうものではない。審査委員の意見を反映させて本論文を見直せば、学術論文としての質はさらに高まるであろう。

以上の所見から、応用事例が限られてきた ASIS による MCMC 法の改良を計量経済分析で広く応用されるパネルデータ回帰モデルや状態空間モデルに拡張したという点において、本論文の学術的貢献は多大であるといえる。したがって、審査委員会は全員一致で本論文が博士（経済学）の学位を授与するにふさわしい内容のものであると判断した。