

Title	データ駆動型アプローチによる高頻度での金融資産価格形成メカニズムの研究
Sub Title	Study on high-frequency price-discovery processes of financial assets in data-driven approach
Author	中妻, 照雄(Nakatsuma, Teruo) 中北, 誠(Nakakita, Makoto) 鳥谷部, 智規(Toyabe, Tomoki)
Publisher	
Publication year	2019
Jtitle	科学研究費補助金研究成果報告書 (2018.)
JaLC DOI	
Abstract	<p>本研究では金融市場における高頻度データ (取引単位で記録されたデータ) の特徴を捉えられるモデルをベイズ推定するための手法の開発に取り組んだ。特に (1) 取引が成立する (約定する) 間隔のモデル化と (2) 短時間における資産収益率の分散のモデル化という2つのテーマに注力した。第1のテーマである約定間隔のモデル化においては、日中季節性をモデルの中で他のパラメータと同時に推定する方法を提案した。一方、第2のテーマである分散のモデル化においても分単位で分散が変動するモデルに同じく日中季節性を導入して他のパラメータと同時に推定する方法を提案した。そして、提案手法の有効性を実際の高頻度データを利用して検証した。</p> <p>In this study, we propose a novel estimation technique for time series models of financial high-frequency data. Specifically, we consider two types of time series models; one is a model of duration between executions of financial transactions while the other is a model of time-varying volatility (variance) in very short intervals. To make these models more realistic, we propose to incorporate intraday seasonality (a cyclical pattern of duration or volatility during trading hours) explicitly into both models and estimate it simultaneously with the model parameters. Since the proposed models are too complex to be estimated with traditional maximum likelihood estimation, we developed an efficient Bayesian Markov chain Monte Carlo (MCMC) method for these models. We applied our new method to real-world high-frequency data (commodity futures and stock prices) and demonstrated their advantage over the conventional models.</p>
Notes	<p>研究種目 : 基盤研究 (C) (一般)</p> <p>研究期間 : 2016 ~ 2018</p> <p>課題番号 : 16K03602</p> <p>研究分野 : ベイズ統計学</p>
Genre	Research Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KAKEN_16K03602seika

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

令和 元 年 6 月 10 日現在

機関番号：32612

研究種目：基盤研究(C)（一般）

研究期間：2016～2018

課題番号：16K03602

研究課題名（和文）データ駆動型アプローチによる高頻度での金融資産価格形成メカニズムの研究

研究課題名（英文）Study on high-frequency price-discovery processes of financial assets in data-driven approach

研究代表者

中妻 照雄（NAKATSUMA, Teruo）

慶應義塾大学・経済学部（三田）・教授

研究者番号：90303049

交付決定額（研究期間全体）：（直接経費） 3,400,000 円

研究成果の概要（和文）：本研究では金融市場における高頻度データ（取引単位で記録されたデータ）の特徴を捉えられるモデルをベイズ推定するための手法の開発に取り組んだ。特に（1）取引が成立する（約定する）間隔のモデル化と（2）短時間における資産収益率の分散のモデル化という2つのテーマに注力した。第1のテーマである約定間隔のモデル化においては、日中季節性をモデルの中で他のパラメータと同時に推定する方法を提案した。一方、第2のテーマである分散のモデル化においても分単位で分散が変動するモデルに同じく日中季節性を導入して他のパラメータと同時に推定する方法を提案した。そして、提案手法の有効性を実際の高頻度データを利用して検証した。

研究成果の学術的意義や社会的意義

近年、金融市場においてミリ秒、マイクロ秒、さらに短い間隔で高速に取引を行って利益を狙うHFT（High-Frequency Trading、高速取引）と呼ばれる手法が急速に普及しており、その影響力が金融市場の安定性を脅かすのではないかと懸念が広がっている。本研究は、より現実的な設定の下で高頻度データの時系列モデルを構築することで、金融市場における資産価格形成メカニズムの理解を深めるとともに、高速取引における新しいリスク管理手法の発展のための一助となることを目指すものである。そして、提案モデルが従来使われてきたモデルよりも現実の高頻度データに対する当てはまりがよいことを示すことに成功した。

研究成果の概要（英文）：In this study, we propose a novel estimation technique for time series models of financial high-frequency data. Specifically, we consider two types of time series models; one is a model of duration between executions of financial transactions while the other is a model of time-varying volatility (variance) in very short intervals. To make these models more realistic, we propose to incorporate intraday seasonality (a cyclical pattern of duration or volatility during trading hours) explicitly into both models and estimate it simultaneously with the model parameters. Since the proposed models are too complex to be estimated with traditional maximum likelihood estimation, we developed an efficient Bayesian Markov chain Monte Carlo (MCMC) method for these models. We applied our new method to real-world high-frequency data (commodity futures and stock prices) and demonstrated their advantage over the conventional models.

研究分野：ベイズ統計学

キーワード：金融高頻度データ 取引間隔 ボラティリティ 日中季節性 板情報 ベイズ推定 マルコフ連鎖モンテカルロ法 モデル選択

1. 研究開始当初の背景

近年、世界各地の金融市場において高速取引 (high frequency trading, HFT) と呼ばれる形態の取引が盛んに行われるようになってきた。名前が示唆するように HFT は高い頻度で株式や外国為替などの売買を繰り返し行う金融取引のことを指す。今では証券市場での取引高の過半を HFT が占めるところまで来ている。

人間には HFT が必要とする極めて短時間での投資判断はできないため、HFT は必然的にコンピュータを駆使したアルゴリズムに基づく取引となる。人間の判断によらない自動取引という特性を持つ HFT の普及は、伝統的ファイナンス理論が想定してきた投資家行動の仮定 (期待効用最大化など) の有効性を弱める可能性がある。事実、HFT のアルゴリズムに依存した取引であるという特性が、経済のファンダメンタルズとは一見無関係な資産価格の急落 (flash crash) という新しい現象を生じさせたと考えられている。コンピュータが次々と高頻度に注文を出し合う HFT が支配的な市場での超短期の値動きのメカニズムは、投資家の合理性を重んじる伝統的ファイナンス理論でも投資家の心理に着目する行動ファイナンス理論でも扱うことが難しい領域といえよう。

高頻度取引市場での価格形成メカニズムは、ファイナンスの理論モデルが依然として十分整備されてはいない領域と考えられる。しかし、HFT の普及の副産物として、研究者にとって利用可能な高頻度の注文データ (ティックデータ) が利用可能となった。ティックデータでは、注文の時間、銘柄、種別 (売りか買い)、数量、約定 (売買取引成立) の有無などが一つのセット (これを板情報という) となって記録されており、そのデータ量は 1 日の取引だけをとっても膨大なものである。この膨大なティックデータを積極的に活用し複雑な現象のモデル化を目指すというデータ駆動型アプローチ (data-driven approach) が高頻度取引市場での価格形成メカニズムの解明を図る上で有用ではないかと考えたのが本研究の発端である。

2. 研究の目的

ティックデータは時間の経過に沿って観測された時系列データの一つであるが、(1) 観測の時間間隔が不規則である、(2) 価格の変化が離散的である、などの特徴があるため、自己回帰移動平均 (ARMA) モデルや一般化自己回帰分散不均一 (GARCH) モデルなどの伝統的に計量ファイナンスで使われてきた時系列モデルの適用は難しい。また、実現分散 (realized variance) のように 1 日の価格変動を単一の記述統計に集約する手法も広く使われるが、これは高頻度での価格形成メカニズムの分析には不向きである。特に詳細な板情報 (約定価格だけではなく売りと買いの注文、価格、数量) を利用できる状況では、過去の約定価格の変動だけではなく、刻々と変化する現時点の注文状況に関する全ての情報を活用できるモデルが好ましいと考えられる。そこで、本研究では板情報の変化を明示的に取り入れて高頻度での価格形成メカニズムをデータ駆動型アプローチでモデル化することを目指す。

しかしながら価格形成メカニズムのモデルを柔軟で現実的なものへと拡張を試みると、自然とモデルは複雑な時系列モデルとなってしまう、伝統的な手法ではモデル内のパラメータの推定が難しくなる。そこで、本研究ではベイズ的アプローチを採用し、マルコフ連鎖モンテカルロ (Markov chain Monte Carlo, MCMC) 法を使ってモデルの推定や検証を行うこととする。そして、MCMC 法で複雑な時系列モデル、特に非線形非正規時系列モデルを効率的にベイズ推定するための新たな手法の確立を本研究におけるもう一つの大きな目的とする。

3. 研究の方法

データ駆動型アプローチによる価格形成メカニズムのモデル化において、本研究では特に (1) 市場で取引が成立する (約定する) 間隔のモデル化と (2) 短時間における資産収益率のボラティリティ (分散) のモデル化という 2 つのテーマに注力する。

第 1 のテーマである約定間隔のモデルとしては、Engle and Russell (1998) による Autoregressive Conditional Duration (ACD) モデルや Bauwens and Veredas (2004) による Stochastic Conditional Duration (SCD) モデルなどが知られている。しかし、これらの約定間隔モデルを推定する際には、日中季節性と呼ばれる高頻度データ特有のパターンに注意しなければならない。ここで日中季節性とは、金融市場における取引の約定間隔が市場の始まった直後と閉まる直前で短くなり、途中の時間帯では反対に長くなる傾向のことを指す。先行研究では約定間隔モデルを推定するための前処理として日中季節性をスプライン回帰などによって除去することが多い。しかし、本研究では日中季節性をスプライン関数で表現して直接 SCD モデルの中に組み込んで他のモデル内のパラメータと同時に MCMC 法でベイズ推定する方法を提案する。

一方、第 2 のテーマであるボラティリティモデルにおいては、短時間で分散が変動する時系列モデルとして Stochastic Volatility (SV) モデルを仮定し、これに先ほどの SCD モデルと同様に日中季節性 (ここで約定間隔の場合とは逆に市場の始まった直後と閉まる直前で分散が大きくなる) のスプライン関数を SV モデルに直接導入して、これを他のパラメータと同時にベイズ推定する方法を提案する。

SCD モデルも SV モデルも非線形非正規状態空間モデルという形式で表現される。一般に状態空間モデルの推定では、観測されない状態変数（SCD モデルでは平均間隔、SV モデルでは分散を変動させる原動力となる）を積分で消去して尤度関数を評価しない限り、最尤法でモデルの推定を行うことは困難である。残念ながら非線形非正規状態空間モデルでは尤度関数の解析的表現は存在しない。そこで本研究では最尤法ではなく MCMC 法によるベイズ的アプローチを用いてモデルの推定を行うことにする。そして、そのための効率的な推定アルゴリズムの開発を目指す。SCD モデルのベイズ推定は Strickland, Forbes and Martin (2006) などで行われているが先行研究の事例は少ない。一方、SV モデルのベイズ推定は、Jacquier, Polson and Rossi (1994) 以来、様々な方法が提案されてきており、高頻度データへの応用事例には Stroud and Johannes (2014) などがある。本研究では、Omori and Watanabe (2008) と Chan and Jeliazkov (2009) の方法を折衷した状態変数に対するブロック・サンプラーを利用することで効率的な MCMC 法のアルゴリズムを構築する。なお SCD モデルと SV モデルは状態空間モデルとして基本的に同じ構造をしているため、よく似た MCMC 法のアルゴリズムを適用できるという利点があることも指摘しておきたい。

最後に本研究の研究体制であるが、研究代表者を統括者とし、研究協力者として2名の大学院生（中北誠と鳥谷部智規）が参加するプロジェクト・チームを編成して研究を遂行した。具体的には、ティックデータの整理（クリーニング）、ティックデータの新しいモデリングの提案、提案モデルの実証のためのコードの開発、提案モデルの検証と応用などの作業を研究代表者と研究協力者で分担し、各人の負担軽減と研究遂行の効率化を図った。

4. 研究成果

まず SCD モデルに関する研究結果について説明する。SCD モデルの推定に対しては、取引間隔の分布に指数分布、ワイブル分布、ガンマ分布、そして対数正規分布の4種類を想定して、MCMC 法による推定アルゴリズムを構築した。対数正規分布の場合には標準的なギブズ・サンプラーが使用できるが、あとの3つの分布に対しては新たに考案した状態変数のブロック・サンプラーと MH アルゴリズムを組み合わせた推定法を適用した。推定アルゴリズムを安定して実行できるプログラムを完成させたのち（本研究では基本的に MATLAB でプログラムを作成した）、商品先物市場（東京商品取引所）と株式市場（東京証券取引所）の高頻度取引データを用いて SCD モデルをベイズ推定し、DIC や WAIC などのモデル選択基準を使って最もふさわしいモデルの定式化の選択を行なった。金融商品や標本期間にもよるが、概ね対数正規分布が選択された。先行研究ではワイブル分布やガンマ分布が取引間隔の分布として使われることが多いが、ベイズ的アプローチにより日中季節性を同時推定することで選択される分布が変わってくることを発見できたのは大きな貢献であると考えられる。

一方、SV モデルの研究では、1分や5分といった短時間における資産価格の収益率を考え、そのボラティリティの立会時間内での変動を、日中季節性を織り込みつつ説明する SV モデルを推定した。この SV モデルにおいては、収益率の分布の裾が厚くなるという傾向を反映させるため、Barndorff-Nielsen (1977) が提案した Generalized Hyperbolic (GH) 分布を仮定した。GH 分布を仮定した日次収益率の SV モデルの研究としては Nakajima and Omori (2012) などがあるが、立会時間内の高頻度収益率に適用した先行研究はあまり知られていない。しかし、GH 分布のままでは一般的すぎるため、その特殊例である t 分布、skew t 分布、ラプラス分布、非対称ラプラス分布、分散ガンマ分布、非対称分散ガンマ分布に限定して分析を行った。さらにレバレッジ効果（収益率と分散の間の相関）もモデルに組み入れ、短時間におけるレバレッジ効果の有無の検証も行った。

ここでも複雑な SV モデルをベイズ推定するために MCMC 法のアルゴリズムを新たに開発している（プログラムは MATLAB で作成した）。そして、この SV モデルに対しても商品先物市場（東京商品取引所）と株式市場（東京証券取引所）の高頻度取引データを用いたベイズ推定を行い、DIC や WAIC などのモデル選択基準を使って最もふさわしいモデルの定式化を選択した。金融商品や標本期間にもよるが、概ねレバレッジ効果は弱く、裾の厚い分布としてラプラス分布が選ばれた傾向が見られた。

以上の研究成果は国内外の学会・研究集会で報告済みであり、学会に参加していた他の研究者からの意見を反映させて論文の修正を行い、近日中に学術誌に投稿する予定である。

参考文献

- Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size, *Proceedings of the Royal Society A*, 353, 401-419.
- Bauwens, L. and Veredas, D. (2004). The stochastic conditional duration model: a latent variable model for the analysis of financial durations, *Journal of Econometrics*, 119, 381-492.
- Chan, J. C. C. and Jeliazkov, I. (2009). Efficient simulation and integrated likelihood estimation in state space models, *International Journal of Mathematical Modelling and Numerical Optimisation*, 1, 101-120.

Engle, R. F. and Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data, *Econometrica*, 66, 1127-1162.

Jacquier, E., Polson, N. G. and Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models, *Journal of Business & Economic Statistics*, 12, 371-389.

Nakajima, J. and Omori, Y. (2012). Stochastic volatility model with leverage and asymmetrically heavy tailed error using GH skew Student's t-distribution. *Computational Statistics & Data Analysis*, 56, 3690-3704.

Omori, Y., and Watanabe, T. (2008). Block sampler and posterior mode estimation for asymmetric stochastic volatility models, *Computational Statistics & Data Analysis*, 52, 2892--2910.

Strickland, C. M., Forbes, C. S. and Martin, G. M. (2006). Bayesian analysis of the stochastic conditional duration model, *Computational Statistics & Data Analysis*, 50, 2247-2267.

Stroud, J. R. and Johannes, M. S. (2014). Bayesian modeling and forecasting of 24-hour high-frequency volatility, *Journal of the American Statistical Association*, 109, 1368-1384.

5 . 主な発表論文等

〔雑誌論文〕(計 1 件)

Taiga Saito, Takanori Adachi, Teruo Nakatsuma, Akihiko Takahashi, Hiroshi Tsuda and Naoyuki Yoshino (2018). Trading and Ordering Patterns of Market Participants in High Frequency Trading Environment: Empirical Study in the Japanese Stock Market, *Asia-Pacific Financial Markets*, 25, 179-220, 査読有
DOI: <https://doi.org/10.1007/s10690-018-9245-6>

〔学会発表〕(計 11 件)

以下のリストでは第一著者が発表者である。

Tomoki Toyabe and Teruo Nakatsuma, Stochastic Conditional Duration Model with Intraday Seasonality and Limit Order Book Information, 12th International Conference on Computational and Financial Econometrics (CFE 2018), 2018 年 12 月 15 日

Teruo Nakatsuma and Makoto Nakakita, Bayesian Analysis of Intraday Stochastic Volatility Models with Skew Heavy-Tailed Error and Smoothing Spline Seasonality, 12th International Conference on Computational and Financial Econometrics (CFE 2018), 2018 年 12 月 15 日

Tomoki Toyabe and Teruo Nakatsuma, Stochastic Conditional Duration Model with Intraday Seasonality and Limit Order Book Information, International Society for Bayesian Analysis (ISBA) World Meeting, 2018 年 6 月 26 日

Sakae Oya and Teruo Nakatsuma, A Graphical Multi-Factor Model of Massively-Many Asset Returns: Application to Long-Term Portfolio Management, International Society for Bayesian Analysis (ISBA) World Meeting, 2018 年 6 月 26 日

Teruo Nakatsuma and Makoto Nakakita, Bayesian Analysis of Intraday Stochastic Volatility Models with Leverage and Skew Heavy-Tailed Error, 11th International Conference on Computational and Financial Econometrics (CFE 2017), 2017 年 12 月 16 日

Ko Sugiura, Teruo Nakatsuma and Akiyoshi Shimura, On Regularized Regression of Categorical Responses on Categorical Predictors, CMStatistics 2017, 2017 年 12 月 16 日

中北 誠・中妻 照雄, Bayesian Analysis of Intraday Stochastic Volatility Models with Leverage and Skew Heavy-Tailed Error in High-Frequency Commodity Market, 統計関連学会連合大会, 2017 年 9 月 4 日

Teruo Nakatsuma and Tomoki Toyabe, Hierarchical Bayes Modeling of Autocorrelation and Intraday Seasonality in Financial Durations, 2nd ISBA-EAC Conference, 2017 年 6 月 6 日

Teruo Nakatsuma and Tomoki Toyabe, Hierarchical Bayes Modeling of Autocorrelation and Intraday Seasonality in Financial Durations, 10th International Conference on Computational and Financial Econometrics (CFE 2016), 2016 年 12 月 9 日

鳥谷部 智規・中妻 照雄, 金先物の取引時間間隔の日中季節性について:階層ベイズモデルによる分析, 統計関連学会連合大会, 2016 年 9 月 5 日

Teruo Nakatsuma and Tomoki Toyabe, Hierarchical Bayes Modeling of Autocorrelation

〔図書〕(計 1 件)

中妻照雄 (2018) 『Python によるファイナンス入門』 朝倉書店 176 ページ .

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕

ホームページ等

<https://github.com/nakatsuma>

6 . 研究組織

(1)研究分担者

なし

(2)研究協力者

研究協力者氏名：中北 誠

ローマ字氏名：(NAKAKITA, Makoto)

研究協力者氏名：鳥谷部 智規

ローマ字氏名：(TOYABE, Tomoki)

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。