

吉田 朋広

東京大学大学院数理科学研究科  
教授

## 先端的確率統計学が開く大規模従属性モデリング

### § 1. 研究成果の概要

連続時間非線形時系列に対するモデリングと解析を可能にする普遍的な数理科学的方法の発見と体系化を試みている。大規模な従属性のあるデータに対するモデリングとその基礎となる確率統計理論の研究、および解析ツールとして役立つソフトウェア開発を行う。高頻度データに基づく統計的モデリング、確率統計理論と計算機統計学的方法の発見、および統計モデリングを可能にする YUIMA 開発を軸に研究が進んだ。

リミットオーダーブック (LOB) は観測間隔が極めて短く従属性の高いビッグデータを与える。その統計的モデリングの過程で得られる方法は普遍的であり、確率過程の統計学に新たな発展をもたらすと期待している。LOB の非定常高頻度従属性データに対して、オーダーの生起確率の評価が可能になる ratio model を提案し、擬似尤度解析によって統計量の性質を導いた。理論研究によって確率過程に対して正当化された情報量規準やスパース推定の方法を用い、オーダーの予測に役立つ共変量の探索を行なった。LOB のモデリングではホークス過程が頻繁に用いられるが、我々が探索した共変量を用いた ratio model はオーダーの予測精度においてそれを凌ぐことを示した。さらに、マーク付きの ratio model への拡張と、推定量の漸近挙動および LOB データへの応用を進めた。

モデリングも予測もできないジャンプで汚染された伊藤過程の観測データの増分の中からジャンプを含む増分を分離し、連続マルチンゲール部分の2次変動過程 (ボラティリティ) を推定する方法を研究した。これまで時間局所的なデータを使った閾値法が提案されてきたが、安定的に良い推定を得ることは難しく、チューニングの方法も確立していない。我々はグローバルジャンプフィルタを導入し、有限時間高頻度観測でボラティリティパラメトリックモデルに対して幾つかのフィルタを提案し、フィルタ付きの擬似最尤推定量および擬似ベイズ推定量の収束率を与え、誤差分布に対する漸近混合正規性と  $L^p$  有界性を研究した。

確率過程に対する統計的モデル選択とスパース推定は大規模従属性モデリングの要であり、 $L^q$

正則化法の理論に関して研究した。Umezumi et al.は非凸ペナルティによる正則化のもとで AIC を構成し、その漸近的性質を擬似尤度解析の方法を使って解明した。この結果は一般的な従属系に対する正則化法の研究に発展している。Suzuki and Yoshida (Japanese Journal of Statistics and Data Science 2020)は、初期推定量を使ったペナルティ付き最小2乗近似法を従属系に一般化し、モデル選択の一致性の誤差確率の精密なオーダー評価を与えた。さらに、最適化が容易な2次形式を使ったモデル選択を行い、その結果に基づいて漸近有効な推定関数によってパラメータの非零成分を2段階推定することでオラクルな推定量を構成した。従属系に対して、通常の初期推定量が適用できることが擬似尤度解析の理論により明らかになっているため、この方法は実用的である。

Kamatani (Stochastic Processes and their Applications 2020)は高次元における random walk Metropolis algorithm の次元依存性を明らかにしたが、さらに、データサイズや次元に対して頑健な手法の開発を行った。また、新しい Piecewise deterministic Monte Carlo 法の研究が進んでいる。

確率微分方程式に対する実用的な統計推測法を開発を行ない、YUIMA への実装を目指している。Nakakita and Uchida (Scandinavian Journal of Statistics 2019)において、観測ノイズ付き高頻度データを用いて、エルゴード的拡散過程モデルのドリフトパラメータと拡散係数パラメータの擬似最尤推定量を導出し、提案した推定量の一致性および漸近正規性を証明した。また、Nakakita and Uchida (Journal of Statistical Planning and Inference 2019)において、観測ノイズ付き高頻度データに基づくエルゴード的拡散過程モデルのドリフトパラメータと拡散係数パラメータの仮説検定法を与えた。観測ノイズを含んだエルゴード的拡散過程モデルに対して、高頻度データを用いて拡散係数パラメータとドリフトパラメータの初期ベイズ推定量および、その初期ベイズ推定量に基づく適応型推定量を導出する統計的手法(ハイブリッド型推測法)を開発した。非定常・非エルゴードな局所安定回帰モデリングにおいて、局所安定型擬似尤度の理論的性質を研究した。とくに、駆動レビ過程の変動指数、モデルのトレンド構造とスケール構造全ての近似信頼集合の構成が可能となった。

確率過程の汎関数の分布の正確な近似が統計推測理論を構築するための基礎となる。Malliavin 解析による漸近分布理論において漸近展開は新たな課題である。Nualart and Yoshida (Electronic Journal of Probability 2019)は Skorohod 積分の漸近展開を非エルゴード系に対して導出したが、応用として、フラクショナルブラウン運動のランダムなウエイト付き2次形式に対する漸近展開を与えた。これは、フラクショナルブラウン運動で記述される証券価格変動のモデリングに繋がる研究である。

### 【代表的な原著論文】

1. Ioane Muni Toke and Nakahiro Yoshida, “Analyzing order flows in limit order books with ratios of Cox-type intensities”, Quantitative Finance, vol. 20, Issue 1, pp.81-98, 2020
2. Yuta Umezu, Yusuke Shimizu, Hiroki Masuda and Yoshiyuki Ninomiya, “AIC for non-concave penalized likelihood method”, Annals of the Institute of Statistical Mathematics, vol. 71, 2, pp.247-274, 2019
3. Shogo H. Nakakita and Masayuki Uchida, “Inference for ergodic diffusions plus noise”, Scandinavian Journal of Statistics, vol. 46, Issue 2, pp.470-516, 2019

## § 2. 研究実施体制

### (1)「東京大学」グループ

① 研究代表者: 吉田 朋広 (東京大学大学院数理科学研究科 教授)

② 研究項目

- ・高頻度時系列データへの統計的モデリングとデータ解析
- ・統計モデリングの基礎となる確率過程の統計学の研究とその応用
- ・確率過程に対する統計解析およびシミュレーションのためのソフトウェア YUIMA III の開発

### (2)「大阪大学」グループ

① 主たる共同研究者: 鎌谷 研吾 (大阪大学大学院基礎工学研究科 准教授)

② 研究項目

- ・確率微分方程式に基づく統計的モデリングと高頻度データ解析
- ・大規模時系列データ解析の基礎となる計算統計理論とその応用
- ・大規模時系列データに有用なモンテカルロ法の開発

### (3)「九州大学」グループ

① 主たる共同研究者: 増田 弘毅 (九州大学大学院数理学研究院 教授)

② 研究項目

- ・非正規レビ過程に関する統計推測の基礎研究とその応用
- ・さまざまな非正規レビ駆動モデルの統計解析ソフトウェアの開発