

論文の内容の要旨

論文題目 Statistical inference for Hawkes processes
(Hawkes過程における統計的推論)

氏 名 郷田昌稔

Hawkes 過程における統計的推論の幾つかの理論を擬似尤度解析に基づき構築し、実際に現実のデータへの応用を行った。本論文は4つの主要部から構成される。

第1部では1次元指数型 Hawkes 過程の高次漸近挙動を取り扱う。具体的には Hawkes 核過程と呼ばれる確率過程を導入し、その汎関数の分布の漸近展開に関する理論を構築した。その重要な応用として、1次元指数型 Hawkes 過程の最尤推定量の誤差分布に対して2次項までの漸近展開を与えた。

第2部では、スパースな構造を持つ多変量指数型 Hawkes 過程を用いて、ウェブサイトにおける複数のグループ間でのテキスト投稿の時間間隔をモデル化する方法を提案する。この際にパラメータの推定方法として擬似最尤推定量と L^1 罰則項付き擬似最尤推定量を組み合わせた方法を導入した。実例として、性別に関連する不満を集めたデータを扱い、年齢と性別でユーザーを12のグループに分類し、各グループ間における不満の伝搬の様子を多変量指数型 Hawkes 過程により定量化した。

第3部では、局外パラメータが存在し、パラメータの真値がパラメータ空間の境界で実現し得る場合に P-O (Penalized Method to Ordinary Method) 推定がオラクル性を満足する十分条件を与え、適当な仮定の下で変数が正しく選択される確率を評価した。その重要な応用として、一般化指数型マーク付き Hawkes 過程の P-O 推定量がオラクル性を満足する事を証明し、変数が正しく選択される確率を評価した。

第4部では、ウェブサイトにおける複数のグループ間でのテキスト投稿の時間間隔の数理モデルとして、Latent Dirichlet Allocation を用いて、文章に含まれる各トピックの割合をマークとして設定した一般化指数型マーク付き Hawkes 過程を提案した。このモデルにより、過去の投稿内容が次の投稿の発生確率に与える影響の定量化を可能とした。実例として、2020年2月1日から2020年9月27日における米国、中国、英国の各大使館のツイッターアカウントによるツイートの時間間隔をモデル化した。解析結果から、Covid-19に関連するツイートが各国のツイートを相互に誘発させている事が確認された。

第1部と第3部は Hawkes 過程の擬似尤度解析に基づき理論が展開される。第2部と第

4 部は Hawkes 過程による実データの解析である。第 3 部の研究は第 2 部の研究から着想を得ている。第 4 部の解析は第 3 部の理論の直接の応用である。