修士論文 要旨

多次元 Hawkes 過程におけるベイズ情報量規準

数理情報学専攻 48206207　岡本 亘

指導教員　荻原 哲平 准教授

1　序論

多次元 Hawkes 過程 [5] は，複数のプロセスが相互に励起するような確率過程であり，SNS における情報拡散や株式市場における注文等の伝播性を持つ現象のモデル化に用いられている．幅広い分野で実証研究が進んでいる多次元 Hawkes 過程であるが，そのモデル選択に関する先行研究は比較的少ない．アプローチは

1. 仮説検定に基づく手法 [3][4]
2. L1 正則化によって変数選択を行う手法 [6]
3. 情報量規準によるスコアリングに基づく手法 [1]

の大きく 3 つに分類できる．1, 2 の手法と比較すると，情報量規準は，実装が容易であるという利点に加え，漸近的一致性などの理論的に望ましい性質が示される．[1] では 1 次元 Hawkes 過程における指数型カーネル関数の個数を選択するという問題設定において, 3 つの情報量規準を適用し，実証的な比較をしている．しかし，多次元 Hawkes 過程において情報量規準を理論的に扱った研究成果はこれまでに発表されていない．そこで本研究では，特にベイズ情報量規準（BIC）に注目し，多次元 Hawkes 過程において BIC が理論的妥当性を持つことと，漸近一致性を満たすことについて示す.

2　多次元 Hawkes 過程

点過程とは，時間軸 $\{t \mid t \in \mathbb{R}_+\}$ 上にランダムに分布する点（イベント）の集合に関する確率過程であり，以下のように定義されるものとする.

定義 2.1 (点過程)**.** 次の 2 つの条件を満たす数直線 $\mathbb{R}_+$ 上の点列 $\{T_n\}_{n\in\mathbb{N}}$ を考える.

1. $0 < T_1 < T_2 < ...$
2. $\lim_{n\to+\infty} T_n = +\infty$

このような点列が，確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上の確率変数としてランダムに定まるとき，

$$N_t(\omega) = \sum_{n\in\mathbb{N}} \mathbf{1}_{\{T_n(\omega)\in[0,t]\}} \tag{1}$$

を点過程といい，発生した点は「イベント」と呼ぶ．イベントの発生履歴に関する情報を，$\mathcal{F}_t = \sigma(\{T_k \mid k \in \mathbb{N} \wedge T_k < t\})$ として保持する.

点過程の振る舞いは強度関数によって特徴づけられる.

定義 2.2 (多次元 Hawkes 過程)**.** 多次元点過程の強度関数が次のように定義されるとき，多次元 Hawkes 過程と呼ぶ．i はプロセスを表す添え字で $i \in \{1, \ldots, d\}$ とする.

$$\lambda^i(t \mid \mathcal{F}_t) = \mu^i + \sum_{j=1}^{d} \int_0^{t-} g^{ij}(t-s)dN_t^j. \tag{2}$$

カーネル関数は以下の通り指数関数型で定義する.

$$g^{ij}(\tau) = \alpha^{ij} \exp(-\beta^{ij}\tau) \qquad (\alpha^{ij} \geq 0, \beta^{ij} > 0). \tag{3}$$

ここで，$\mathbf{\Phi} = (\frac{\alpha^{ij}}{\beta^{ij}})_{1\leq i,j\leq d}$ という行列を定義し，その最大固有値 $\rho(\mathbf{\Phi})$ が 1 未満となるように α^{ij}, β^{ij} の範囲を制限する.

3　ベイズ情報量規準と先行研究

$m = 1, \cdots, M$ に対して，ベイズモデル

$$\mathcal{M}_m = \{((\mathfrak{p}_m, \pi_{m,n}(\boldsymbol{\theta}_m), l_{m,n}(\cdot \mid \boldsymbol{\theta}_m)) \mid \boldsymbol{\theta}_m \in \Theta_m\} \tag{4}$$

を考える．ただし，$\mathfrak{p}_m$ はモデル $\mathcal{M}_m$ の事前選択確率で $\sum_{m=1}^M \mathfrak{p}_m = 1$, $\pi_{m,n} : \Theta_m \to (0, \infty)$ は事前分布, $\Theta_m \subset \mathbb{R}^{p_m}$: 有界領域，$l_{m,n}(\cdot \mid \boldsymbol{\theta}_m)$ はモデル $\mathcal{M}_m$ の対数尤度関数，p_m はパラメータ空間 Θ_m の次元である．$l_{m,n}(\cdot \mid \boldsymbol{\theta}_m)$ はパラメータ空間の閉包 $\overline{\Theta}_m$ 上で連続であると仮定する．このとき，データ $\boldsymbol{X}_n$ が与えられたときの最尤推定量を $\hat{\boldsymbol{\theta}}_{m,n} = \underset{\boldsymbol{\theta}_m\in\overline{\Theta}_m}{\operatorname{argmax}} l_{m,n}(\boldsymbol{X}_n \mid \boldsymbol{\theta}_m)$ と定義する．最尤推定量は次の収束レート $\sqrt{n}$ によりあるパラメータ $\boldsymbol{\theta}_{m,0} \in \Theta_m$ に収束するとする.

定義 3.1 (モデルの事後選択確率)**.** ベイズの定理から，モデルの事後選択確率が定まる.

$$P_m := \frac{\mathfrak{p}_m \int_{\Theta_m} \exp\{l_{m,n}(x \mid \boldsymbol{\theta}_m)\}\pi_{m,n}(\boldsymbol{\theta}_m)d\boldsymbol{\theta}_m}{\sum_{i=1}^M \mathfrak{p}_i \int_{\Theta_i} \exp\{l_{i,n}(x \mid \boldsymbol{\theta}_i)\}\pi_{i,n}(\boldsymbol{\theta}_i)d\boldsymbol{\theta}_i}. \tag{5}$$

[2] は一般的な確率過程において，対数尤度関数の $\boldsymbol{\theta}_{m,0}$ 周りでの二次のテイラー展開の誤差の確率収束に関する仮定（Assumption 3.1 in [2]），事前分布に関する

標準的な仮定（Assumption 3.2 in [2]），対数尤度関数の裾の挙動を制御する仮定（Assumption 3.3 in [2]）等の条件の下で，BIC が定数オーダーを除いて，事後選択確率を最大化することを示した．

4 本研究の主要な結果

定義 2.2 において相互励起のパラメータ α^{ij} に零制約を付す場合とそうでない場合について場合分けを考えると，全部で 2^{d^2} パターンのモデルが考えられる．これを $\mathcal{M}_m$ とする．また，真のパラメータを $\boldsymbol{\theta}_0$ とし，パラメータ空間 Θ_m に $\boldsymbol{\theta}_0$ を含むモデルのうち，Θ_m の次元が一番小さいものを真のモデル $\mathcal{M}_{m_0}$ と定義する．

仮定 4.1. 事前分布 $\pi_{m,n}$ は次の 2 つを満たす．

1. 任意の $n \in \mathbb{N}$ に対して，$\pi_{m,n}(\boldsymbol{\theta}_{m,0}) > 0$ が成り立ち，$\sup_n \sup_{\boldsymbol{\theta}_m} \pi_{m,n}(\boldsymbol{\theta}_m) < +\infty$ となる．
2. 任意の $K > 0$ に対して，$\lim_{n\to+\infty} \sup_{|\boldsymbol{u}|<K} |\pi_{m,n}(\boldsymbol{\theta}_{m,0} + n^{-1/2}u) - \pi_{m,n}(\boldsymbol{\theta}_{m,0})| = 0$.

定理 4.2 (多次元 Hawkes 過程におけるベイズ情報量規準)**.** 事前分布 $\pi_{m,n}(\boldsymbol{\theta}_m)$ が仮定 4.1 を満たし，$\Theta_{m_0} \subset \Theta_m$ のとき，

$$\mathrm{BIC}_{m,n} = -2l_{m,n}(\hat{\boldsymbol{\theta}}_{m,n}) + p_m \log n \tag{6}$$

と定義すると，以下が成り立つ．

$$\log P_m = -\frac{1}{2}\mathrm{BIC}_{m,n} + O_p(1) + C_n. \tag{7}$$

ただし，C_n はモデルに依存しない共通項である．

これにより，BIC を最小化するモデル $\mathcal{M}_m$ は定数オーダーを除いて，事後選択確率を最大化することが分かる．

定理 4.3 (多次元 Hawkes 過程における BIC の一致性)**.** $m \in \{1, \cdots, M\}\backslash\{m_0\}$ として，モデル $\mathcal{M}_m$ の事前分布 $\pi_{m,n}(\boldsymbol{\theta}_m)$ が仮定 4.1 を満たすとする．このとき，BIC は真のモデル $\mathcal{M}_{m_0}$ への一致性を満たす．

$$\lim_{n\to+\infty} \mathbb{P}(\mathrm{BIC}_{m_0,n} - \mathrm{BIC}_{m,n} < 0) = 1. \tag{8}$$

紙面の制約上，証明に関しては本論文に譲る．

5 数値実験

2 次元 Hawkes 過程について相互励起の有無を場合分けすることで，16 個のモデルが定まる．これを用いて，情報量規準が真のモデルを選択するか調べるための数値実験を実施した．実験の手順は以下の通りである．

1. 真のパラメータ $\boldsymbol{\theta}_0$ と対応する真のモデル $\mathcal{M}_{m_0}$ を用意する．$\boldsymbol{\theta}_0$ の下で多次元 Hawkes 過程に従うデータを観察期間 $[0, T]$ の間でサンプリングする．
2. 得られた人工データについて，16 個の部分モデルすべてを候補モデルとして，最尤推定を行い，$\mathrm{AIC}_m, \mathrm{BIC}_m$ をそれぞれ計算する．
3. $\mathrm{AIC}_m, \mathrm{BIC}_m$ を最小化するモデル $\mathcal{M}_{m_1}$ を求める．
4. 1 ~ 3 を 1000 回繰り返す．$\mathcal{M}_{m_1}$ と $\mathcal{M}_{m_0}$ が一致する回数を記録し，正解率を計算する．

この実験を $T = 1000, 2000, 4000, 8000, 16000, 32000$ の 6 ケースについて行い，観察期間を長くすることでモデル選択の正解率がどのように推移するかを調べた．また 16 個の部分モデルすべてについて真のモデルとして設定して実験を行った．

紙面の制約上，結果の詳細を載せることはできないが，すべてのモデルについて，T を長くすることで，BIC が真のモデルを正しく選択できることを確認した．

6 結論

本研究では，多次元 Hawkes 過程において BIC が定数オーダーを除いてベイズのモデル事後選択確率を最大化すること，および漸近一致性を満たすことを示した．さらに，人工データを用いた数値実験により AIC・BIC を用いたモデル選択を行い，特に BIC においてデータ数が多くなるにつれて真のモデルが正しく選択されることを確認した．今後の課題としては，実データを用いた他のモデル選択手法との比較等が考えられる．

参考文献

[1] Chen, J., Hawkes, A., Scalas, E. and Trinh, M.: Performance of information criteria for selection of Hawkes process models of financial data, *Quantitative Finance*, Vol. 18, No. 2, pp. 225–235 (2018).

[2] Eguchi, S. and Masuda, H.: Schwarz type model comparison for LAQ models, *Bernoulli*, Vol. 24, No. 3, pp. 2278 – 2327 (2018).

[3] Eichler, M., Dahlhaus, R. and Dueck, J.: Graphical modeling for multivariate hawkes processes with nonparametric link functions, *Journal of Time Series Analysis*, Vol. 38, No. 2, pp. 225–242 (2017).

[4] Embrechts, P. and Kirchner, M.: Hawkes graphs, *Theory of Probability & Its Applications*, Vol. 62, No. 1, pp. 132–156 (2018).

[5] Hawkes, A. G.: Spectra of some self-exciting and mutually exciting point processes, *Biometrika*, Vol. 58, No. 1, pp. 83–90 (1971).

[6] Xu, H., Farajtabar, M. and Zha, H.: Learning granger causality for hawkes processes, *International Conference on Machine Learning*, PMLR, pp. 1717–1726 (2016).