

時空間データのロバスト統計

RA 清智也

情報理工学系研究科数理情報学専攻

概要

本研究の目的は、時間的・空間的なデータに対するロバストな統計的推測手法を確立することである。特に、有界領域で密にサンプリングを行なった場合に現れる漸近理論 (fixed domain asymptotics) に着目し、推定やモデル選択に関する諸結果を得た。

キーワード: 局所漸近混合正規性, fixed domain asymptotics, フラクタル指数, 変換ガウスモデル。

1 変換ガウスモデルの LAMN 性

ガウス過程をある変換関数で変換して得られる確率過程のモデル (変換ガウスモデルと呼ぶ) を想定し、変換関数の形を観測データから適応的に求めたい状況を考える (図 1)。このような場合、比較的多くのパラメータを持つ関数群を用意し、AIC などのモデル選択規準を使うことが考えられる。しかし、AIC を導出する際的前提である局所漸近正規性 (LAN) が、変換ガウスモデルについて成り立つかどうかを確認する必要がある。これに対して本研究では、変換ガウスモデルは LAN でなく、局所漸近混合正規性 (LAMN) と呼ばれる性質を持つことを示した [2, 3]。LAMN モデルのためのモデル選択規準は既に昨年研究成果として得られている [1]。したがって、変換ガウスモデルのためのモデル選択規準が構築できたことになる。

ここで LAMN 性について概説しておく。 θ をモデルに含まれる未知のパラメータとし、モデルの尤度を $L_n(\theta)$ とおく (n はデータ数)。LAMN 性の本質的な要請は、対数尤度の二回微分を規格化した量 $J_n = -n^{-1}(\log L_n)''$ が、 $n \rightarrow \infty$ のとき適

当な確率変数 J に収束することである。この確率変数 J はランダムフィッシャー情報量と呼ばれる。変換ガウスモデルの一例について、 J_n の分布関数が収束する様子を図 2 に示す。

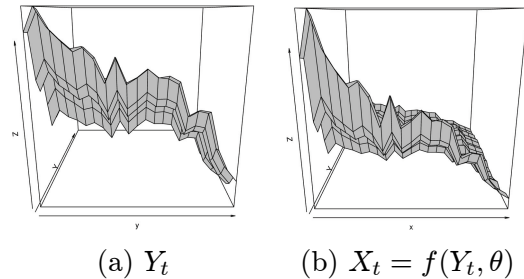


図 1: 2 径数変換ガウスモデル。 ($Y_t \mid t \in [0, 1]^2$) はガウス過程に従い、有限個の t に対して $X_t = f(Y_t, \theta)$ が観測されるものとする。 f は既知の関数、 θ は未知のパラメータである。

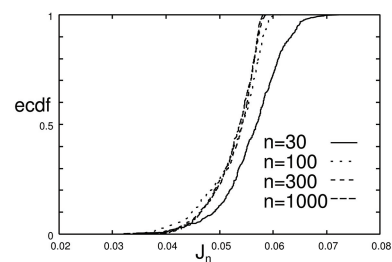


図 2: ランダムフィッシャー情報量への収束の様子。

2 フラクタル指数の推定問題

前節で述べた変換ガウスモデルの LAMN 性は、その証明において少し強い仮定を置いていた。それは、変換前のガウス過程が独立増分性を持つ、と

いうものである。この仮定は空間データ解析の立場からは必ずしも現実的ではなく、より弱い仮定で考察する必要があった。そこでまず、1 径数ガウス過程の fixed domain asymptotics における振る舞いを調べることにした。同じ設定下において、M. L. Stein [5] はフラクタル指数の擬似最尤推定量 (QMLE) を提案した。本研究では、この推定量の漸近的性質を示すことに成功した。

以下、問題設定と結果を簡単に述べる。 $\{X_t \mid t \in [0, 1]\}$ を実数値定常ガウス過程とし、その自己相関関数 $\rho(t)$ の形を次のように仮定する：

$$\rho(t) = 1 - \nu|t|^\alpha + o(|t|^\alpha) \quad \text{as } t \rightarrow 0. \quad (1)$$

このとき α はフラクタル指数と呼ばれ、標本 X_t の滑らかさを指定する重要な量となる (図 3)。観測データ $\{X_{i/n}\}_{i=1}^n$ から α を推定する推定量 $\hat{\alpha}$ として性能の良いものを見つきたい。Stein は、データの差分に対してスペクトル密度を用いる推定量 QMLE を提案し、精度が良いことを数値的に示した。QMLE は、差分データが、(1) 式の誤差項に関わらず fractional Gaussian noise と呼ばれる特別な時系列に収束することを利用している。本研究では QMLE の漸近正規性・最適性を証明し、またバイアス項のオーダーを決定した。図 4 に、既存の推定量と比較した数値実験結果の一例を示す。

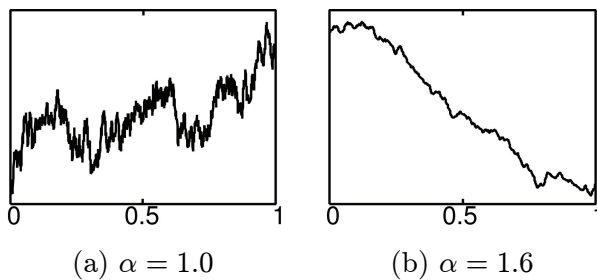


図 3: フラクタル指数と標本過程。

3 おわりに

1 節の結果は雑誌に投稿中である [2]。また国際学会において発表し、好評を得た [3]。2 節の内容は、学位論文に含まれており [4]、近く雑誌に投稿する予定である。

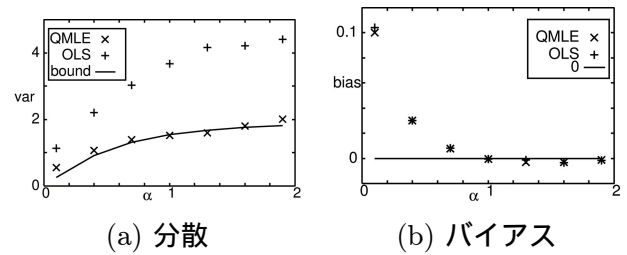


図 4: QMLE と Increment-based 推定量 (OLS) の比較 (Matérn モデル, $n = 2048$)。横軸は真の α を表す。

参考文献

- [1] Sei, T. and Komaki, F. (2004). Bayesian prediction and model selection for locally asymptotically mixed normal models. Tech. Rep. METR2004-25, The University of Tokyo, Submitted for publication.
- [2] Sei, T. (2004a). Local asymptotic mixed normality of transformed Gaussian models for random fields. Tech. Rep. METR2004-37, The University of Tokyo, Submitted for publication.
- [3] Sei, T. (2004b). An asymptotic property of transformed Gaussian models for random fields. Proceedings of International Conference on Interdisciplinary Mathematical and Statistical Techniques, Dec. 27-29, 2004, SMC, India.
- [4] Sei, T. (2005). Asymptotic properties of estimators and information criteria for random fields. Ph. D. thesis, The University of Tokyo.
- [5] Stein, M. L. (1995). Fixed domain asymptotics for spatial periodograms. *J. Amer. Statist. Assoc.*, **90**, 1277-1288.