

# Lévy 駆動型確率微分方程式の段階的推定について

上原 悠禎<sup>†</sup>・増田 弘毅<sup>†</sup>

(受付 2016 年 6 月 28 日；改訂 10 月 13 日；採択 10 月 17 日)

## 要 旨

非正規型 Lévy 過程で駆動される確率微分方程式モデルの推定を考える。指数的エルゴード性とデータの高頻度性の下、正規型疑似スコア関数に基づいてスケール係数およびドリフト係数をこの順に段階的に推定する方法を提案し、推定量の漸近正規性および裾確率評価を導出する。推定対象を分割することで最適化の計算負荷が削減され、より安定した推定精度が得られる。拡散過程の場合と異なり、特に両係数のパラメータが共通因子を持つ場合には、ドリフト係数の漸近共分散行列は同時推定の場合と異なる形をとる。

キーワード：エルゴード性、正規型疑似スコア関数、高頻度観測、段階的推定、Lévy 駆動型確率微分方程式。

## 1. はじめに

連続時間確率過程で記述される統計モデルによって観測頻度に伴う推定精度の差異がモデリング可能となり、モデルの具体系のみならず推定方式によって漸近分布論も様々な形式を取り得る。本稿では、常微分方程式  $dx_t = a(x_t, \alpha)dt$  の時間発展に確率的なノイズを加えた、フィルター付き確率空間  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, P)$  上に定義された次の  $d$  次元確率微分方程式 (SDE) モデル

$$(1.1) \quad dX_t = a(X_t, \alpha, \gamma)dt + c(X_{t-}, \gamma)dJ_t.$$

を考える。ここで  $J$  は  $(\mathcal{F}_t)$ -適度な  $r$  次元非正規型 Lévy 過程であり、ドリフト係数  $a: \mathbb{R}^d \times \Theta_\alpha \times \Theta_\gamma \rightarrow \mathbb{R}^d$  とスケール係数  $c: \mathbb{R}^d \times \Theta_\gamma \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$  は有限次元未知パラメータ

$$\theta := (\alpha, \gamma) \in \Theta := \Theta_\alpha \times \Theta_\gamma \subset \mathbb{R}^{p_\alpha} \times \mathbb{R}^{p_\gamma}$$

を除いて既知であるとし、初期変数  $X_0$  は  $J$  と独立とする。簡単のためパラメータ空間  $\Theta$  は有界凸領域とする。SDE モデル (1.1) の解過程から高頻度観測  $(X_{t_0}, X_{t_1}, \dots, X_{t_n})$  が得られていると仮定する。ここで、 $t_j = t_j^n = jh_n$  であり、正数列  $(h_n)_{n \in \mathbb{N}}$  は

$$(1.2) \quad T_n := nh_n \rightarrow \infty, \quad nh_n^2 \rightarrow 0$$

を満たしているとする。目的は  $\theta$  の真値  $\theta_0 \in \Theta$  を推定することである。

Lévy 過程とは独立定常増分性を有する確率過程であり、離散時間ランダムウォークの自然な連続時間版と解釈される。これは互いに独立な Wiener 過程と非正規型 (純粋ジャンプ型) Lévy 過程の畳み込みとして、また同時に複合 Poisson 過程のある種の極限として定義される。そのジャンプ構造・挙動の多様性により、Lévy 過程は非正規性が色濃い金融データや工学・生体信

<sup>†</sup>九州大学大学院 数理学研究院：〒 819-0395 福岡県福岡市西区元岡 744

号の表現として広く用いられている。Lévy 過程の理論を包括的に扱った標準的な教科書として Sato (1999) を挙げておく。

現在膨大な高頻度データの入手が可能となり、上述の現象記述精度の高さから Lévy 駆動型 SDE モデルの実現象への応用の機運が高まっている。しかしながらその高頻度統計理論は  $J$  に対応する無限分解可能分布族を (相当) 限定した場合でさえ未だ十分に整備されていない。これは一つには、ジャンプ構造が過度に多様なため統一的な推定手法の構築が不可能であるという事実が災いしている。実際、最尤推定量の最適収束率も観測頻度とジャンプ構造の双方が本質的に関与して決まる。この問題への一つの対処法として、Masuda (2013) は (1.1) でドリフト係数  $a(x, \alpha, \gamma)$  が  $\gamma$  によらない場合に、漸近効率を犠牲にした正規型疑似最尤推定量の漸近挙動を調べた。この手法により、漸近有効性を放棄することで陽な疑似尤度関数を用いつつ駆動ノイズ  $J$  の具体的な分布構造を仮定せずに  $\theta$  の推定を行える。

ところで単一の疑似尤度に基づいた  $\theta$  の推定は  $\alpha$  と  $\gamma$  の同時最適化を伴い、Masuda (2013) の推定方式も例外ではない。このため、係数が複雑な場合には計算負荷の高さおよび同時最適化の不安定さが懸念されよう。近年、拡散過程 ( $J$  が標準 Wiener 過程) の場合にこの問題に対する種々のアプローチ (Uchida and Yoshida, 2012 や Kamatani and Uchida, 2015 とその参考文献を参照) が提案され、拡散係数とドリフト係数を個別に、しかも漸近有効性を保持しつつ推定する形で、数値統計の観点からも発展がなされた。これらの方法論においては  $\alpha$  と  $\gamma$  の最適収束率の差異が有効活用されている。翻って Lévy 駆動型 SDE の場合についてはその限りではない。実際、正規型疑似尤度は  $\alpha, \gamma$  共に  $\sqrt{T_n}$  の収束率しか持たず、さらに一般に漸近共分散行列は拡散過程の場合のようにブロック対角型にはならないため、段階的な推定手法の定式化は自明な話ではない。この点を踏まえ、本稿の主眼は  $\theta$  の推定を  $\alpha, \gamma$  個々の推定へ分割する段階的推定法を定式化し、当該推定量の漸近分布と多項式型裾確率評価を導出することにある。提案手法によって  $\gamma, \alpha$  をこの順に別個に推定可能となり、計算負荷の大きな軽減の可能性が期待できる。

本稿の構成は以下の通りである。まず 2 節で段階的推定量  $\hat{\theta}_n$  構成の概略を述べ、その漸近的性質を導出するための仮定を導入する。3 節において主結果である  $\hat{\theta}_n$  の漸近正規性および裾確率評価を与える。特に、漸近分布の形を通じて拡散過程の場合との差異を浮き彫りにする。4 節では段階的推定量のパフォーマンスを確認するために幾つかの数値実験結果を紹介し、最後に 5 節において主結果の証明を与える。

## 2. 段階的係数推定法

本稿の主眼である (1.1) の段階的係数推定法について述べていく。

$J$  が有限共分散をもつとき、標準化した  $\tilde{J}_t := \text{cov}(J_1)^{-1/2} \{J_t - E(J_1)t\}$  を用いて (1.1) を書き換えると

$$(2.1) \quad dX_t = (a(X_t, \alpha, \gamma) + c(X_t, \gamma)E(J_1))dt + c(X_{t-}, \gamma)\text{cov}(J_1)^{1/2}d\tilde{J}_t$$

となる。新たに  $E(J_1), \text{cov}(J_1)$  をパラメータとみなせば (2.1) は (1.1) での特別な場合となるため、(1.1) において一般性を失うことなく  $J$  は標準化されているとしてよい。ここではより強く、以下を仮定する ( $I_r$  は  $r$  次元単位行列を表す)。

**仮定 2.1.**  $E(J_1) = 0, E(J_1^{\otimes 2}) = I_r$ , また任意の  $q > 0$  に対し  $E[|J_1|^q] < \infty$ .

任意の確率過程  $Y$  と  $\mathbb{R}^d \times \bar{\Theta}$  上の可測関数  $f$  に対して

$$Y_j = Y_{t_j}, \quad \Delta_j Y = Y_j - Y_{j-1}, \quad f_j(\theta) = f(Y_j, \theta)$$

と適宜略記する. (1.1)の形式的な Euler-丸山近似は

$$(2.2) \quad X_j \approx X_{t_{j-1}} + h_n a_{j-1}(\alpha, \gamma) + c_{j-1}(\gamma) \Delta_j J$$

で与えられる. 以下では

$$S := c^{\otimes 2}$$

と表す. ただし任意の行列に対し  $x^{\otimes 2} := xx^T$  ( $T$  は転置を表す) である. 形式的に局所正規近似  $\Delta_j J \approx N(0, h_n I_r)$  を適用した場合の条件付き尤度の近似

$$(2.3) \quad \mathcal{L}(X_j | X_{t_{j-1}} = x) \approx N(x + h_n a(x, \alpha, \gamma), h_n S(x, \gamma))$$

に基づき, Masuda (2013)では, ドリフト係数とスケール係数内のパラメータの重複がない場合(ドリフト係数が  $a(x, \alpha)$  の形)において, 同時正規型疑似尤度

$$(2.4) \quad \mathbb{M}_n(\theta) := -\frac{1}{2} \sum_{j=1}^n \left\{ \log |S_{j-1}(\gamma)| + \frac{1}{h_n} S_{j-1}^{-1}(\gamma) [(\Delta_j X - h a_{j-1}(\theta))^{\otimes 2}] \right\}$$

から得られる疑似最尤推定量の漸近正規性, モーメント収束を導出した. 局所正規近似(2.3)そのものは理論上正しくないが, 条件付き平均・共分散構造に関するモーメント法の一種として推定に利用可能であることが分かる. このような正規型疑似尤度の有用性は時系列モデルにおいてもよく知られているが, Lévy 駆動型 SDE モデルの場合には収束率が真に下がる.

以下では  $\partial_x$  で変数  $x$  に関する偏微分を, また  $|A|$  で正方行列  $A$  の行列式を表す. また, 多重線形形式  $M := \{M^{(i_1 \dots i_K)} : i_k = 1, \dots, d_k; k = 1, \dots, K\} \in \mathbb{R}^{d_1} \otimes \dots \otimes \mathbb{R}^{d_K}$  と  $d_k$  次元ベクトル  $u_k = \{u_k^{(j)}\}$  について

$$M[u_1, \dots, u_K] := \sum_{i_1=1}^{d_1} \dots \sum_{i_K=1}^{d_K} M^{(i_1, \dots, i_K)} u_1^{(i_1)} \dots u_K^{(i_K)}$$

と表す.

正規型疑似スコア関数  $\theta \mapsto (\mathbb{G}_{1,n}(\gamma), \mathbb{G}_{2,n}(\alpha, \gamma))$ , ただし

$$\begin{aligned} \mathbb{G}_{1,n}(\gamma) &:= \frac{1}{n} \sum_{j=1}^n \left\{ \partial_\gamma \log |S_{j-1}(\gamma)| + \frac{1}{h_n} \partial_\gamma (S_{j-1}^{-1})(\gamma) [(\Delta_j X)^{\otimes 2}] \right\} \\ &= \frac{1}{n} \sum_{j=1}^n \left\{ [\text{trace}(S_{j-1}^{-1} \partial_\gamma S_{j-1})(\gamma)]_{i=1}^{p_\gamma} - \frac{1}{h_n} (S_{j-1}^{-1}(\partial_\gamma S_{j-1}) S_{j-1}^{-1})(\gamma) [(\Delta_j X)^{\otimes 2}] \right\}, \end{aligned}$$

$$\mathbb{G}_{2,n}(\alpha, \gamma) := \frac{1}{T_n} \sum_{j=1}^n S_{j-1}^{-1}(\gamma) [\partial_\alpha a_{j-1}(\theta), \Delta_j X - h_n a_{j-1}(\theta)],$$

に対して以下で定義される段階的推定量  $\hat{\theta}_n := (\hat{\alpha}_n, \hat{\gamma}_n)$  を提案する.

- (1)  $\hat{\gamma}_n$  を  $\mathbb{G}_{1,n}(\gamma) = 0$  のランダムな解とする.
- (2) 上記  $\hat{\gamma}_n$  を用いて,  $\hat{\alpha}_n$  を  $\mathbb{G}_{2,n}(\alpha, \hat{\gamma}_n) = 0$  のランダムな解とする.

このように定義された  $\hat{\theta}_n$  は, 元々の疑似尤度  $\mathbb{M}_n$  による  $(p_\alpha + p_\gamma)$ -次元同時最適化を  $\gamma$  と  $\alpha$  個別の最適化へ分断可能とするため, 推定値計算負荷の削減による時間短縮へつながら.

$\mathbb{G}_{1,n}(\gamma)$  は,  $\mathbb{M}_n$  の定義でドリフト項を無いものと勝手にみなした場合の  $\gamma$  に関する疑似スコア関数をスケールしたものに相当する. このドリフト非依存正規型疑似尤度は拡散過程の拡散項の推定において今日標準的となったが, そこに至るまでは当該分野において数多の試行錯誤があったと思われる. 近年においては, 計算負荷の観点から疑似尤度関数の同時最適化変

数の削減や初期推定量の構築などの目的にも利用されている．関連する先行研究については，Uchida and Yoshida (2012) や Kamatani and Uchida (2015) とその参考文献を参照されたい．

ドリフト項，拡散項に共通パラメータを持つ拡散過程に上記の段階的推定法を適用すると，収束レートの違い(ドリフトパラメータは  $\sqrt{T_n}$ ，拡散パラメータは  $\sqrt{n}$ ) から，その漸近分散は共通パラメータを持たない場合と同様のものが得られる．しかしながら，Lévy 駆動型 SDE において正規型疑似尤度最尤推定量はドリフトパラメータ，スケールパラメータともに収束レートが  $\sqrt{T_n}$  であるため(Masuda, 2013)，状況は本質的に異なる．実際，後述の定理 3.4 によって，疑似スコア関数  $(\mathbb{G}_{1,n}(\gamma), \mathbb{G}_{2,n}(\alpha, \gamma))$  に基づく段階的推定量  $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\gamma}_n)$  は，ドリフト・スケール係数が共通パラメータを持つ場合には  $\hat{\alpha}_n$  の漸近分散が変わることが明示的に示される．

### 3. 主結果

本節の定理 3.4 において，推定量の漸近正規性に加え，推定量依存型統計量の確率・モーメント評価を保証する裾確率評価を与える．

**仮定 3.1.** ドリフト係数  $a(\cdot, \theta_0)$  とスケール係数  $c(\cdot, \gamma_0)$  は Lipschitz 連続で，各  $i \in \{0, 1, 2\}$ ， $j \in \{0, \dots, 5\}$ ， $k \in \{0, \dots, 5\}$  に対して連続な導関数  $\partial_x^i \partial_\alpha^j \partial_\gamma^k a$  および  $\partial_x^i \partial_\gamma^k c$  が存在し，ある非負定数  $C_{(i,j,k)}$  に対して

$$(3.1) \quad \sup_{(x, \alpha, \gamma) \in \mathbb{R}^d \times \bar{\Theta}_\alpha \times \bar{\Theta}_\gamma} \frac{1}{1 + |x|^{C_{(i,j,k)}}} \{|\partial_x^i \partial_\alpha^j \partial_\gamma^k a(x, \alpha, \gamma)| + |\partial_x^i \partial_\gamma^k c(x, \gamma)| + |S^{-1}(x, \gamma)|\} < \infty$$

が成り立つ．

**仮定 3.2.** ある確率測度  $\pi_0$  と正定数  $a, C$  に対して

$$(3.2) \quad \sup_{f: |f| \leq g_q} \left| \int f(y) \{P(X_t \in dy | X_0 = x) - \pi_0(dy)\} \right| \leq C e^{-at} g_q(x), \quad x \in \mathbb{R}, q > 0,$$

が成立する ( $g_q(x) := 1 + |x|^q$ )．さらに任意の  $q > 0$  に対して

$$(3.3) \quad \sup_{t \in \mathbb{R}_+} E[|X_t|^q] < \infty.$$

仮定 3.2 の下，エルゴード定理より任意の多項式増大関数  $f$  について

$$\frac{1}{n} \sum_{j=1}^n f(X_{t_{j-1}}) \xrightarrow{P} \int f(x) \pi_0(dx)$$

が成立する．さらに大数の法則の収束率を加味したモーメント評価を得ることができる(Masuda, 2013, Lemma 4.3)．

関数  $\mathbb{G}^\infty(\theta) := (\mathbb{G}_\alpha^\infty(\alpha), \mathbb{G}_\gamma^\infty(\gamma)) \in \mathbb{R}^p$  および定数行列  $\mathcal{I}_\alpha \in \mathbb{R}_{p_\alpha} \otimes \mathbb{R}_{p_\alpha}$ ， $\mathcal{I}_\gamma \in \mathbb{R}_{p_\gamma} \otimes \mathbb{R}_{p_\gamma}$  を以下で定義する:  $u, u' \in \mathbb{R}^{p_\alpha}$  と  $v, v' \in \mathbb{R}^{p_\gamma}$  に対して

$$\mathbb{G}_\alpha^\infty(\alpha) := \int \partial_\alpha a(x, \theta_0)^T S^{-1}(x, \gamma_0) [a(x, \alpha_0, \gamma_0) - a(x, \alpha, \gamma_0)] \pi_0(dx),$$

$$\mathbb{G}_\gamma^\infty(\gamma) := \int \partial_\gamma S^{-1}(x, \gamma) [S(x, \gamma_0) - S(x, \gamma)] \pi_0(dx),$$

$$\mathcal{I}_\alpha[u, u'] := - \int (S^{-1}(x, \gamma_0) [\partial_\alpha a(x, \theta_0)[u], \partial_\alpha a(x, \theta_0)[u']]) \pi_0(dx),$$

$$\mathcal{I}_\gamma[v, v'] := \int \text{trace}\{(S^{-1} \partial_\gamma S \otimes S^{-1} \partial_\gamma S)(x, \gamma_0)[v, v']\} \pi_0(dx).$$

モデルの識別可能性・非負定値性のため、以下の仮定を導入する。

仮定 3.3.  $\mathcal{I}_\alpha, \mathcal{I}_\gamma$  はそれぞれ逆行列を持ち、さらにある正定数  $\chi_\alpha, \chi_\gamma$  が存在して

$$(3.4) \quad |\mathbb{G}_\alpha^\infty(\alpha)| \geq \chi_\alpha |\alpha_0 - \alpha|, \quad |\mathbb{G}_\gamma^\infty(\gamma)| \geq \chi_\gamma |\gamma - \gamma_0|,$$

が任意の  $(\alpha, \gamma)$  に対して成り立つ。

$\nu_0$  で  $J$  の Lévy 測度を表すとき、 $i_1, \dots, i_m \in \{1, \dots, r\}$  に対して

$$\nu_{i_1, \dots, i_m}(m) = \int z^{(i_1)} \dots z^{(i_m)} \nu_0(dz)$$

と書く(仮定 2.1 の下、これらは有限値として定まる)。記号

$$\begin{aligned} \Sigma &= \begin{pmatrix} \Sigma_\alpha & \Sigma_{\alpha, \gamma} \\ \text{Sym.} & \Sigma_\gamma \end{pmatrix}, \quad \hat{\Sigma}_n = \begin{pmatrix} \hat{\Sigma}_{\alpha, n} & \hat{\Sigma}_{\alpha, \gamma, n} \\ \text{Sym.} & \hat{\Sigma}_{\gamma, n} \end{pmatrix}, \\ \mathcal{I} &:= \begin{pmatrix} \mathcal{I}_\alpha & B \\ 0 & \mathcal{I}_\gamma \end{pmatrix}, \quad \hat{\mathcal{I}}_n := \begin{pmatrix} -\hat{\Sigma}_{\alpha, n} & \hat{B}_n \\ 0 & \hat{\mathcal{I}}_{\gamma, n} \end{pmatrix} \end{aligned}$$

を導入する。ここで  $u, u' \in \mathbb{R}^{p_\alpha}$  と  $v, v' \in \mathbb{R}^{p_\gamma}$  に対して、

$$\begin{aligned} \Sigma_\alpha[u, u'] &:= \int S^{-1}(x, \gamma_0) [\partial_\alpha a(x, \theta_0)[u], \partial_\alpha a(x, \theta_0)[u']] \pi_0(dx), \\ \Sigma_\gamma[v, v'] &:= \int \sum_{s, t, s', t'=1}^r \nu_{s, t, s', t'}(4) \{ \partial_\gamma S^{-1}(x, \gamma_0) [v, c^{(\cdot s)}(x, \gamma_0), c^{(\cdot t)}(x, \gamma_0)] \} \\ &\quad \cdot \{ \partial_\gamma S^{-1}(x, \gamma_0) [v', c^{(\cdot s')}(x, \gamma_0), c^{(\cdot t')}(x, \gamma_0)] \} \pi_0(dx), \\ \Sigma_{\alpha, \gamma}[u, v] &:= \int \sum_{k_1, k_2, k_3=1}^r \nu_{k_1, k_2, k_3}(3) S^{-1}(x, \gamma_0) [\partial_\alpha a(x, \theta_0)[u], c^{(\cdot k_1)}(x, \gamma_0)] \\ &\quad \cdot \partial_\gamma S^{-1}(x, \gamma_0) [v, c^{(\cdot k_2)}(x, \gamma_0), c^{(\cdot k_3)}(x, \gamma_0)] \pi_0(dx), \\ B[u, v] &:= - \int S^{-1}(x, \gamma_0) [\partial_\alpha a(x, \theta_0)[u], \partial_\gamma a(x, \theta_0)[v]] \pi_0(dx), \\ \hat{\Sigma}_{\alpha, n}[u, u'] &:= \frac{1}{n} \sum_{j=1}^n S_{j-1}^{-1}(\hat{\gamma}_n) [\partial_\alpha a_{j-1}(\hat{\theta}_n)[u], \partial_\alpha a_{j-1}(\hat{\theta}_n)[u']], \\ \hat{\Sigma}_{\gamma, n}[v, v'] &:= \frac{1}{T_n} \sum_{j=1}^n (\partial_\gamma S^{-1} \otimes \partial_\gamma S^{-1})(\hat{\gamma}_n) [(v, (\Delta_j X)^{\otimes 2}), (v', (\Delta_j X)^{\otimes 2})], \\ \hat{\Sigma}_{\alpha, \gamma, n}[u, v] &:= \frac{1}{T_n} \sum_{j=1}^n (S_{j-1}^{-1} \otimes \partial_\gamma S_{j-1}^{-1})(\hat{\gamma}_n) \\ &\quad \cdot [(\partial_\alpha a_{j-1}(\hat{\theta}_n)[u], \Delta_j X - h_n a_{j-1}(\hat{\theta}_n)), (v, (\Delta_j X)^{\otimes 2})], \\ \hat{B}_n[u, v] &:= - \frac{1}{n} \sum_{j=1}^n S_{j-1}^{-1}(\hat{\gamma}_n) [\partial_\alpha a_{j-1}(\hat{\theta}_n)[u], \partial_\gamma a_{j-1}(\hat{\theta}_n)[v]], \\ \hat{\mathcal{I}}_{\gamma, n}[v, v'] &:= \frac{1}{n} \sum_{j=1}^n \text{trace}\{ (S_{j-1}^{-1} \partial_\gamma S_{j-1} \otimes S_{j-1}^{-1} \partial_\gamma S_{j-1})(\hat{\gamma}_n)[v, v'] \} \end{aligned}$$

である。最後に  $\phi(\cdot; 0, V)$  で  $N(0, V)$  の密度関数を表す。

**定理 3.4.** 仮定 2.1, 3.1, 3.2, 3.3, さらに行列  $\Sigma$  の正定値性の下で以下が成り立つ.

(1) 裾確率評価: 任意の  $K > 0$  に対してある定数  $C_K > 0$  が存在して

$$(3.5) \quad \sup_n P(|\sqrt{T_n}(\hat{\alpha}_n - \alpha_0)| > r) + \sup_n P(|\sqrt{T_n}(\hat{\gamma}_n - \gamma_0)| > r) \leq \frac{C_K}{r^K}, \quad r > 0.$$

(2) 漸近標準正規性:

$$(3.6) \quad \sqrt{T_n} \hat{\Sigma}_n^{-1/2} \hat{\mathcal{L}}_n(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N(0, I_p).$$

(3) モーメント収束: 高々多項式増大度を持つ任意の連続関数  $f$  に対して

$$(3.7) \quad E[f(\sqrt{T_n}(\hat{\theta}_n - \theta_0))] \rightarrow \int f(y)\phi(y; 0, \mathcal{I}^{-1}\Sigma(\mathcal{I}^{-1})^\top)dy.$$

**注意 3.5.** 定理 3.4 より特に, 正規型疑似尤度を適用する場合の  $\Delta_j X$  の微小時間変動においてはスケール係数部分が支配的であることがわかる. すなわち, 元々の  $\alpha, \gamma$  の同時推定量の収束率は共に同じ  $\sqrt{T_n}$  でしかもその同時推定量は漸近的に直交する(独立である)とは限らない状況であるにも関わらず, ドリフト係数を完全に無視しても,  $\gamma$  の同時推定と段階的推定は漸近同等となる.

**注意 3.6.** 行列  $B$  は共通パラメータの存在により生じる有限バイアスの補正的役割を果たしている. 実際, 係数間のパラメータの重複がない場合には  $B = 0$  であり,  $\hat{\alpha}_n$  と  $\hat{\gamma}_n$  の漸近分散は Masuda (2013) のそれと等しく, 推定分離による推定量の分散の変化は生じない. さらにこの時,  $\nu(3) = 0$  であれば  $\hat{\alpha}_n$  と  $\hat{\gamma}_n$  は漸近的に直交する(独立になる).

**注意 3.7.**  $d = r$  ( $\dim(X_t) = \dim(J_t)$ ) の場合には, 駆動ノイズの増分の代替として Euler 残差  $\hat{\delta}_j := c_{j-1}(\hat{\gamma}_n)^{-1}(\Delta_j X - h_n a_{j-1}(\hat{\theta}_n))$  を用いることで, Masuda and Uehara (2017) と同様, 駆動ノイズ  $J$  に関するモーメント法を定式化できる. 証明の概略, 推定関数の満たすべき条件等は同論文を参照されたい.

**例 3.8.** (2.1) の部分族である一次元 SDE モデル

$$(3.8) \quad dX_t = \left( a_0(X_t, \gamma) + \sum_{l=1}^{p_\alpha} \alpha_l a_{1l}(X_t, \gamma) \right) dt + \exp\left(\frac{1}{2} \sum_{k=1}^{p_\gamma} \gamma_k c_k(X_{t-})\right) dJ_t$$

は, 係数の状態変数  $X_t$  について多様な非線形性を表現可能であると同時に,  $\gamma$  および  $\alpha$  の推定の各段階が凸最適化になるという点において特殊である. ここで

$$a_0(x, \gamma), \quad a_1(x, \gamma) = (a_{11}(x, \gamma), \dots, a_{1p_\alpha}(x, \gamma)), \quad c(x) = (c_1(x), \dots, c_{p_\gamma}(x)),$$

はその関数形が既知であるとしている. このとき,  $\hat{\gamma}_n$  は推定方程式

$$\sum_{j=1}^n \left( 1 - \frac{(\Delta_j X)^2}{h_n} \exp(-\gamma \cdot c(X_{j-1})) \right) c(X_{j-1}) = 0$$

の解であり,  $\hat{\alpha}_n$  は

$$\hat{\alpha}_n = \frac{1}{h_n} \left( \sum_{j=1}^n \exp(-\hat{\gamma}_n \cdot c(X_{j-1})) a_1^{\otimes 2}(X_{j-1}, \hat{\gamma}_n) \right)^{-1} \cdot \sum_{j=1}^n \exp(-\hat{\gamma}_n \cdot c(X_{j-1})) \{ \Delta_j X - h_n a_0(X_{j-1}, \hat{\gamma}_n) \} a_1(X_{j-1}, \hat{\gamma}_n)$$

と陽に与えられる。また、他の第一段階が凸最適化となる scale 係数の関数形の例として、 $c(x, \gamma) = (\sum_{k=1}^{p\gamma} \gamma_k c_k(x))^{-1/2}$  などが挙げられる。

#### 4. 数値実験

先に述べた通り、駆動ノイズは基準化されたものを考えればよい。まず次の SDE モデルを取り扱う。

$$(4.1) \quad dX_t = \left\{ \frac{\gamma X_t}{1 + X_t^2} - \left( \alpha_1 X_t + \frac{\alpha_2}{(1 + X_t^2)^\gamma} \right) \right\} dt + c(X_{t-}, \gamma) dJ_t, \quad X_0 = 0.$$

ここで、スケール係数および各パラメータの真値は以下の 2 種類を想定する。

$$(i) \ c(x, \gamma) = \exp\left(-\frac{\gamma}{1+x^2}\right) \text{ および } \theta_0 = (\alpha_{1,0}, \alpha_{2,0}, \gamma_0) = (1, 2, 3)$$

$$(ii) \ c(x, \gamma) = -\frac{\gamma}{1+x^2} \text{ および } \theta_0 = (\alpha_{1,0}, \alpha_{2,0}, \gamma_0) = (1, 2, 1)$$

また、 $J_1$  の分布は、 $NIG(1, 0, t, 0)$ ,  $NIG(10, 0, 10t, 0)$ ,  $NIG(25/3, 20/3, 9/5t, -12/5t)$  の 3 通りを対象とする。 $NIG$  分布は  $IG$  分布の正規平均尺度混合により定義される確率分布 (Barndorff-Nielsen, 1998) であり、 $NIG(\alpha, \beta, \delta, \mu)$  の特性関数は、

$$\phi_{X_1}(u) = \exp(-\delta(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2})) \exp(i\mu u),$$

で与えられる。この表現と Lévy-Kchintchin の公式から、

$$E[X_t] = t \left( \mu + \delta \frac{\beta}{\sqrt{\alpha^2 - \beta^2}} \right), \quad \text{Var}[X_t] = t\delta \frac{\alpha^2}{(\sqrt{\alpha^2 - \beta^2})^3}$$

が得られる。これらを用いて、 $E[J_t] = 0, E[J_t^2] = t$  であることが確認できる。また、Masuda (2013, Proposition 5.4) と係数の形を考慮すれば、(4.1) は仮定 3.2 を満たしていることがわかる。

本数値実験では、各駆動ノイズについて  $(T_n, h_n, n)$  を  $(10, 0.05, 200)$ ,  $(10, 0.01, 1000)$ ,  $(50, 0.05, 1000)$ ,  $(50, 0.01, 5000)$ ,  $(100, 0.05, 2000)$ ,  $(100, 0.01, 10000)$  の 6 通りに分け、それぞれ 500 回ずつ独立に Euler-丸山法に基づきパスを生成し段階的推定量  $\hat{\theta}_n := (\hat{\alpha}_{1,n}, \hat{\alpha}_{2,n}, \hat{\gamma}_n)$  を計算した。(i) の  $\gamma$  についてはパラメータ空間は  $\Theta_\gamma = [0, 10]$  とし、初期値は 2 として最適化を行った(他の推定量は陽に書ける)。また、パスの生成と  $\hat{\gamma}_n$  の最適化には R の YUIMA パッケージ (Brouste et al., 2014) を用いた。表 1-3 にそれらの推定結果を示した。

- ドリフトパラメータの推定結果を見ると、表 1-3 全ての場合においてターミナル  $T_n$  の増加に応じて推定精度が向上している。また、ターミナル固定の下で観測幅が小さくなくても推定精度の向上はほとんど見られない。これは  $\hat{\alpha}_n$  の収束レートが  $\sqrt{T_n}$  であること、また注意 3.5 で触れた通り  $X$  の微小時間変動においてはスケール項が支配的であることから妥当な結果だと考えられる。

- スケールパラメータの推定精度は表 1-3 を見るとターミナル  $T_n$  の増加に応じて良くなっており、これは拡散過程とは異なる傾向であるが、ここでは  $\hat{\gamma}_n$  の収束レートが  $\sqrt{T_n}$  であることから自然である。観測幅  $h_n$  の観点からは、表 1, 表 3 ではターミナル固定の下で観測幅が小さくなる時、標準偏差の改善が見られない。しかし表 2 では、ターミナル固定の下では観測幅の小さい場合の推定精度が大きい場合の推定精度を優越している。この点は、 $NIG(\delta, 0, \delta, 0)$  は  $\delta \rightarrow \infty$  で  $N(0, 1)$  に全変動収束するため、ノイズの微小時間増分系列について正規近似が機能していると捉えれば整合的である。

- スケール係数の違いに関して見ると、表 1-3 いずれにおいても (i) より (ii) の推定精度が良

表 1.  $\mathcal{L}(J_t) = NIG(1, 0, t, 0)$  の場合の段階的推定量  $\hat{\theta}_n := (\hat{\alpha}_{1,n}, \hat{\alpha}_{2,n}, \hat{\gamma}_n)$  の標準平均・標準標準偏差(括弧内).

| $T_n$ | $n$   | $h_n$ | (i)                  |                      |                  | (ii)                 |                      |                  |
|-------|-------|-------|----------------------|----------------------|------------------|----------------------|----------------------|------------------|
|       |       |       | $\hat{\alpha}_{1,n}$ | $\hat{\alpha}_{2,n}$ | $\hat{\gamma}_n$ | $\hat{\alpha}_{1,n}$ | $\hat{\alpha}_{2,n}$ | $\hat{\gamma}_n$ |
| 10    | 200   | 0.05  | 0.97                 | 2.09                 | 2.65             | 1.12                 | 2.31                 | 0.95             |
|       |       |       | (0.20)               | (0.60)               | (0.73)           | (0.26)               | (0.80)               | (0.23)           |
| 10    | 1000  | 0.01  | 1.08                 | 2.06                 | 3.04             | 1.14                 | 2.38                 | 0.97             |
|       |       |       | (0.26)               | (0.64)               | (1.01)           | (0.26)               | (0.88)               | (0.23)           |
| 50    | 1000  | 0.05  | 0.98                 | 2.00                 | 2.84             | 1.00                 | 2.00                 | 0.98             |
|       |       |       | (0.14)               | (0.51)               | (0.56)           | (0.11)               | (0.39)               | (0.11)           |
| 50    | 5000  | 0.01  | 1.04                 | 2.01                 | 3.07             | 1.03                 | 2.07                 | 1.00             |
|       |       |       | (0.17)               | (0.41)               | (0.62)           | (0.12)               | (0.41)               | (0.13)           |
| 100   | 2000  | 0.05  | 0.98                 | 1.99                 | 2.88             | 0.98                 | 1.94                 | 1.00             |
|       |       |       | (0.11)               | (0.35)               | (0.40)           | (0.08)               | (0.27)               | (0.10)           |
| 100   | 10000 | 0.01  | 1.01                 | 2.03                 | 3.01             | 1.01                 | 2.03                 | 1.00             |
|       |       |       | (0.12)               | (0.34)               | (0.44)           | (0.08)               | (0.27)               | (0.10)           |

表 2.  $\mathcal{L}(J_t) = NIG(10, 0, 10t, 0)$  の場合の段階的推定量  $\hat{\theta}_n := (\hat{\alpha}_{1,n}, \hat{\alpha}_{2,n}, \hat{\gamma}_n)$  の標準平均・標準標準偏差(括弧内).

| $T_n$ | $n$   | $h_n$ | (i)                  |                      |                  | (ii)                 |                      |                  |
|-------|-------|-------|----------------------|----------------------|------------------|----------------------|----------------------|------------------|
|       |       |       | $\hat{\alpha}_{1,n}$ | $\hat{\alpha}_{2,n}$ | $\hat{\gamma}_n$ | $\hat{\alpha}_{1,n}$ | $\hat{\alpha}_{2,n}$ | $\hat{\gamma}_n$ |
| 10    | 200   | 0.05  | 0.92                 | 2.15                 | 2.55             | 1.11                 | 2.31                 | 0.99             |
|       |       |       | (0.16)               | (0.55)               | (0.22)           | (0.27)               | (0.87)               | (0.06)           |
| 10    | 1000  | 0.01  | 0.99                 | 2.17                 | 2.88             | 1.14                 | 2.44                 | 1.00             |
|       |       |       | (0.15)               | (0.59)               | (0.17)           | (0.27)               | (0.88)               | (0.04)           |
| 50    | 1000  | 0.05  | 0.97                 | 2.07                 | 2.88             | 1.00                 | 2.00                 | 0.99             |
|       |       |       | (0.07)               | (0.48)               | (0.13)           | (0.12)               | (0.43)               | (0.03)           |
| 50    | 5000  | 0.01  | 1.00                 | 2.13                 | 2.97             | 1.02                 | 2.06                 | 1.00             |
|       |       |       | (0.07)               | (0.53)               | (0.09)           | (0.12)               | (0.42)               | (0.02)           |
| 100   | 2000  | 0.05  | 0.99                 | 1.98                 | 2.95             | 0.99                 | 1.97                 | 0.99             |
|       |       |       | (0.05)               | (0.45)               | (0.10)           | (0.09)               | (0.31)               | (0.02)           |
| 100   | 10000 | 0.01  | 1.00                 | 2.08                 | 2.99             | 1.01                 | 2.04                 | 1.00             |
|       |       |       | (0.05)               | (0.48)               | (0.06)           | (0.09)               | (0.32)               | (0.01)           |

いことが確かめられるが、これは当然パラメータの真値に依存して結果は異なる。 $\hat{\gamma}_n$  の漸近分散に着目すると、(ii) の設定(パラメータが線形に含まれている場合)だと漸近分散が不変測度  $\pi_0$  の変化に加えて、真値の 2 乗に比例することが容易に確認できる(ケース (i) では  $\hat{\gamma}_n$  の漸近分散は  $\pi_0(dy)$  を通じてのみ  $\theta_0$  に依存する)。ここには記載していないが、実際に (ii) の場合に真値を大きくすると標準偏差が悪化することを確認している。

また、図 1, 2 において  $\mathcal{L}(J_1) = NIG(25/3, 20/3, 9/5t, -12/5t)$ ,  $(T_n, n, h_n) = (100, 10000, 0.01)$  の場合(表 3) のヒストグラムおよび標準正規密度関数(実線)を示した。標準正規分布近似は概ねうまくいっている。

次に scale 項に複数のパラメータを持つ例として下記の SDE モデルを想定する:

$$(4.2) \quad dX_t = \left( -\alpha_1 X_t + \frac{\alpha_2}{1 + X_t^2} \right) dt + \exp \left( -\frac{\gamma_0 + \gamma_1 X_t}{2(1 + X_t^2)} \right) dJ_t, \quad X_0 = 0.$$

ここで、 $\mathcal{L}(J_t) = NIG(10, 0, 10t, 0)$ ,  $\Theta_\gamma = [0, 50] \times [0, 50]$ ,  $\theta_0 = (\alpha_{1,0}, \alpha_{2,0}, \gamma_{1,0}, \gamma_{2,0}) = (1, 2, 3, 4)$ ,



表 3.  $\mathcal{L}(J_t) = NIG(25/3, 20/3, 9/5t, -12/5t)$  の場合の段階的推定量  $\hat{\theta}_n := (\hat{\alpha}_{1,n}, \hat{\alpha}_{2,n}, \hat{\gamma}_n)$  の標準平均・標準標準偏差(括弧内).

| $T_n$ | $n$   | $h_n$ | (i)                  |                      |                  | (ii)                 |                      |                  |
|-------|-------|-------|----------------------|----------------------|------------------|----------------------|----------------------|------------------|
|       |       |       | $\hat{\alpha}_{1,n}$ | $\hat{\alpha}_{2,n}$ | $\hat{\gamma}_n$ | $\hat{\alpha}_{1,n}$ | $\hat{\alpha}_{2,n}$ | $\hat{\gamma}_n$ |
| 10    | 200   | 0.05  | 0.93<br>(0.13)       | 2.13<br>(0.42)       | 2.56<br>(0.49)   | 1.08<br>(0.24)       | 2.26<br>(0.84)       | 0.98<br>(0.17)   |
| 10    | 1000  | 0.01  | 1.03<br>(0.14)       | 2.11<br>(0.45)       | 2.99<br>(0.67)   | 1.13<br>(0.24)       | 2.45<br>(0.90)       | 1.00<br>(0.17)   |
| 50    | 1000  | 0.05  | 0.95<br>(0.07)       | 2.02<br>(0.33)       | 2.79<br>(0.36)   | 0.99<br>(0.10)       | 1.97<br>(0.37)       | 0.99<br>(0.09)   |
| 50    | 5000  | 0.01  | 1.01<br>(0.08)       | 2.06<br>(0.34)       | 3.04<br>(0.40)   | 1.02<br>(0.10)       | 2.08<br>(0.37)       | 1.00<br>(0.07)   |
| 100   | 2000  | 0.05  | 0.96<br>(0.06)       | 1.97<br>(0.27)       | 2.84<br>(0.28)   | 0.98<br>(0.07)       | 1.94<br>(0.26)       | 0.99<br>(0.05)   |
| 100   | 10000 | 0.01  | 1.00<br>(0.06)       | 2.01<br>(0.29)       | 2.98<br>(0.29)   | 1.01<br>(0.07)       | 2.04<br>(0.28)       | 1.00<br>(0.05)   |

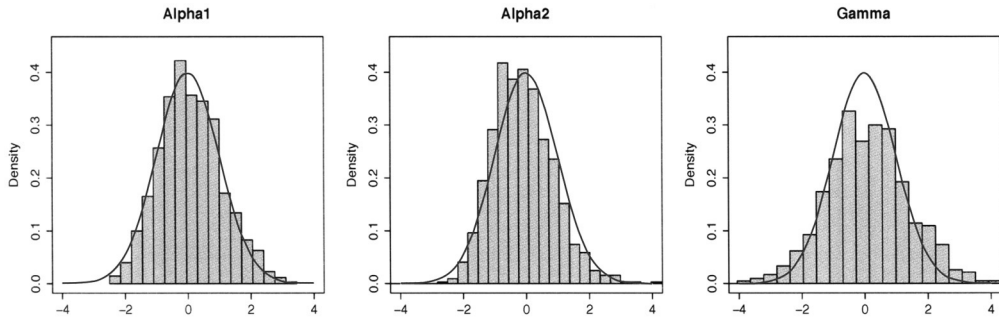


図 1. 標準化(スチューデント化)された段階的推定量  $\hat{\theta}_n$  のヒストグラム (i) の場合).

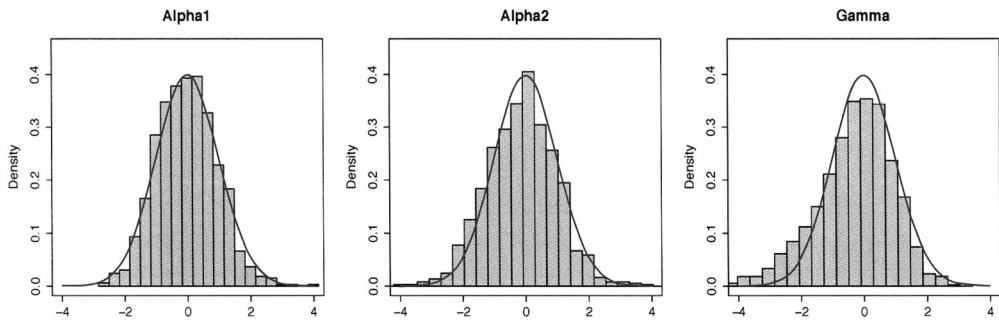


図 2. 標準化(スチューデント化)された段階的推定量  $\hat{\theta}_n$  のヒストグラム (ii) の場合).

初期値をパラメータ空間上の一様乱数とし, その他の設定は先程のものと同様とする. 数値実験結果を表 4 に示す. ここでは比較のため拡散過程  $(\mathcal{L}(J_t) = N(0, t))$  とした場合の実験結果も併記した. 表 4 から, 両モデルともにドリフトの推定精度は比較的良い一方で, scale パラメータの推定値の平均が真値とかけ離れており, 今のモデルでは平均構造を無視したことで生じるバイアスの影響が顕著であると考えられる. この問題に対し, 三段階目の推定として通常の正規

表 4.  $\mathcal{L}(J_t) = NIG(10, 0, 10t, 0)$  および  $\mathcal{L}(J_t) = N(0, t)$  の場合の段階的推定量  $\hat{\theta}_n := (\hat{\alpha}_{1,n}, \hat{\alpha}_{2,n}, \hat{\gamma}_{1,n}, \hat{\gamma}_{2,n})$  の標準平均・標本標準偏差(括弧内).

| $T_n$ | $n$   | $h_n$ | $NIG(10, 0, 10t, 0)$ |                      |                      |                      | $\mathcal{L}(J_t) = N(0, t)$ |                      |                      |                      |
|-------|-------|-------|----------------------|----------------------|----------------------|----------------------|------------------------------|----------------------|----------------------|----------------------|
|       |       |       | $\hat{\alpha}_{1,n}$ | $\hat{\alpha}_{2,n}$ | $\hat{\gamma}_{1,n}$ | $\hat{\gamma}_{2,n}$ | $\hat{\alpha}_{1,n}$         | $\hat{\alpha}_{2,n}$ | $\hat{\gamma}_{1,n}$ | $\hat{\gamma}_{2,n}$ |
| 10    | 200   | 0.05  | 1.02<br>(0.20)       | 2.03<br>(0.34)       | 0.61<br>(0.42)       | 6.48<br>(0.54)       | 1.02<br>(0.19)               | 2.02<br>(0.33)       | 0.64<br>(0.42)       | 6.45<br>(0.53)       |
| 10    | 1000  | 0.01  | 1.04<br>(0.17)       | 2.07<br>(0.29)       | 2.08<br>(0.37)       | 4.94<br>(0.44)       | 1.04<br>(0.17)               | 2.07<br>(0.29)       | 2.08<br>(0.37)       | 4.94<br>(0.44)       |
| 50    | 1000  | 0.05  | 0.97<br>(0.12)       | 1.94<br>(0.22)       | 1.07<br>(0.40)       | 6.02<br>(0.44)       | 0.97<br>(0.12)               | 1.94<br>(0.22)       | 1.12<br>(0.37)       | 5.98<br>(0.40)       |
| 50    | 5000  | 0.01  | 1.01<br>(0.11)       | 2.02<br>(0.21)       | 2.28<br>(0.31)       | 4.74<br>(0.33)       | 1.01<br>(0.11)               | 2.02<br>(0.21)       | 2.32<br>(0.24)       | 4.70<br>(0.25)       |
| 100   | 2000  | 0.05  | 0.96<br>(0.08)       | 1.92<br>(0.16)       | 1.39<br>(0.38)       | 5.70<br>(0.40)       | 0.97<br>(0.09)               | 1.93<br>(0.17)       | 1.41<br>(0.35)       | 5.68<br>(0.37)       |
| 100   | 10000 | 0.01  | 1.00<br>(0.08)       | 2.00<br>(0.16)       | 2.43<br>(0.29)       | 4.58<br>(0.30)       | 1.01<br>(0.09)               | 2.01<br>(0.17)       | 2.46<br>(0.21)       | 4.56<br>(0.22)       |

表 5.  $\mathcal{L}(J_t) = NIG(10, 0, 10t, 0)$  および  $\mathcal{L}(J_t) = N(0, t)$  の場合の適応型推定法に基づく推定量  $\tilde{\gamma}_n := (\tilde{\gamma}_{1,n}, \tilde{\gamma}_{2,n})$  の標準平均・標本標準偏差(括弧内).

| $T_n$ | $n$   | $h_n$ | $NIG(10, 0, 10t, 0)$   |                        | $\mathcal{L}(J_t) = N(0, t)$ |                        |
|-------|-------|-------|------------------------|------------------------|------------------------------|------------------------|
|       |       |       | $\tilde{\gamma}_{1,n}$ | $\tilde{\gamma}_{2,n}$ | $\tilde{\gamma}_{1,n}$       | $\tilde{\gamma}_{2,n}$ |
| 10    | 200   | 0.05  | 3.17<br>(0.82)         | 4.05<br>(0.89)         | 3.29<br>(0.73)               | 3.92<br>(0.79)         |
| 10    | 1000  | 0.01  | 3.06<br>(0.50)         | 3.99<br>(0.55)         | 3.06<br>(0.50)               | 3.99<br>(0.55)         |
| 50    | 1000  | 0.05  | 3.08<br>(0.65)         | 4.12<br>(0.67)         | 3.20<br>(0.60)               | 4.00<br>(0.63)         |
| 50    | 5000  | 0.01  | 3.00<br>(0.41)         | 4.04<br>(0.42)         | 3.06<br>(0.29)               | 3.98<br>(0.30)         |
| 100   | 2000  | 0.05  | 3.05<br>(0.60)         | 4.14<br>(0.61)         | 3.13<br>(0.55)               | 4.06<br>(0.56)         |
| 100   | 10000 | 0.01  | 3.00<br>(0.36)         | 4.04<br>(0.37)         | 3.04<br>(0.25)               | 4.00<br>(0.26)         |

型疑似尤度(2.4)に第二段階で得られた  $\hat{\alpha}_n$  をプラグインし, scale パラメータの推定値をアップデートすることを考える. すなわち,  $\gamma$  の推定値を以下の関数の最大点として新たに定義する:

$$\gamma \mapsto -\frac{1}{2} \sum_{j=1}^n \left\{ \log |S_{j-1}(\gamma)| + \frac{1}{h_n} S_{j-1}^{-1}(\gamma) [(\Delta_j X - h a_{j-1}(\hat{\alpha}_n, \gamma))^{\otimes 2}] \right\}.$$

ここで最適化の初期値には第一段階で得た推定量を用いることとする. このアイデアは拡散過程の適応型推定法(cf. Uchida and Yoshida, 2012)に基づくものである. 観測頻度条件(1.2)の下, 本稿で扱っている Lévy 型 SDE の scale 推定に関しても, 三段階目の推定量の漸近挙動は一段階目の推定量のそれと等しいことが示される. 表 5 にアップデートした推定量  $\tilde{\gamma}_n$  を示した. 拡散過程と同様に, 推定量の平均値に関する改善が見て取れる. 標準偏差は悪化しているが, これは上述の通り漸近分散の改善が理論上見込めないことや,  $\hat{\alpha}_n$  の推定値の揺らぎに起因するものであると考えられる.

### 5. 定理 3.4 の証明

まず本節で用いる記号を述べておく． $E^{j-1}[\cdot]$  で  $\mathcal{F}_{t_{j-1}}$  に関する条件付き期待値を表す．任意の  $\mathbb{R} \times \Theta$  上の関数  $f$  について  $f_s(\theta) = f(X_s, \theta)$  と定義する (記号  $f_{j-1}(\theta) = f(X_{t_{j-1}}, \theta)$  との重複はあるが, 混乱は生じないであろう)．特に  $\theta = \theta_0$  の時は  $f_s = f(X_s, \theta_0)$  と略記する．また

$$\begin{aligned} M'(x, \theta) &:= \partial_\alpha a(x, \theta)^T S^{-1}(x, \gamma) \in \mathbb{R}^{p_\alpha} \otimes \mathbb{R}^d, \\ M''(x, \gamma) &:= -\partial_\gamma S^{-1}(x, \gamma) \in \mathbb{R}^{p_\gamma} \otimes \mathbb{R}^d \otimes \mathbb{R}^d \end{aligned}$$

とする．ある正定数  $C$  が存在して十分大きな任意の  $n$  について  $x_n \leq Cy_n$  となる時  $x_n \lesssim y_n$  と書く ( $C$  は  $n$  に依存せず, 現れるごとに異り得る)．

#### 5.1 裾確率評価 (3.5) の証明

$(\hat{\gamma}_n, \hat{\alpha}_n)$  は, コントラスト関数  $\mathbb{M}_{1,n}(\gamma) := -nh_n |\mathbb{G}_{1,n}(\gamma)|^2$ ,  $\mathbb{M}_{2,n}(\alpha) := -nh_n |\mathbb{G}_{2,n}(\alpha, \hat{\gamma}_n)|^2$  に対応した  $M$  推定量とみなすことができる．これらに対して以下の確率場を定義する:

$$\begin{aligned} \mathbb{J}_{1,n}(u_1) &:= \exp \left\{ \mathbb{M}_{1,n} \left( \gamma_0 + \frac{1}{\sqrt{T_n}} u_1 \right) - \mathbb{M}_{1,n}(\gamma_0) \right\}, \\ \mathbb{J}_{2,n}(u_2) &:= \exp \left\{ \mathbb{M}_{2,n} \left( \alpha_0 + \frac{1}{\sqrt{T_n}} u_2, \hat{\gamma}_n \right) - \mathbb{M}_{2,n}(\alpha_0, \hat{\gamma}_n) \right\}, \end{aligned}$$

ここで  $u_1, u_2$  はそれぞれ, 集合  $U_{1,n} := \{u_1 \in \mathbb{R}^{p_\gamma} : \gamma_0 + \frac{1}{\sqrt{T_n}} u_1 \in \Theta_\gamma\}$ ,  $U_{2,n} := \{u_2 \in \mathbb{R}^{p_\alpha} : \alpha_0 + \frac{1}{\sqrt{T_n}} u_2 \in \Theta_\alpha\}$  の元とする． $\mathbb{J}_{1,n}(u_1), \mathbb{J}_{2,n}(u_2)$  の定義から,

$$\sqrt{T_n}(\hat{\gamma}_n - \gamma_0) \in \operatorname{argmax}_{u_1 \in U_{1,n}} \mathbb{J}_{1,n}(u_1), \quad \sqrt{T_n}(\hat{\alpha}_n - \alpha_0) \in \operatorname{argmax}_{u_2 \in U_{2,n}} \mathbb{J}_{2,n}(u_2)$$

である．ゆえに, 任意の  $r > 0$  について

$$P(|\sqrt{T_n}(\hat{\gamma}_n - \gamma_0)| > r) \leq P \left( \sup_{u_1: |u_1| > r, u_1 \in U_{1,n}} \mathbb{J}_{1,n}(u_1) \geq \mathbb{J}_{1,n}(0) \equiv 1 \right)$$

が成り立つ ( $\sqrt{T_n}(\hat{\alpha}_n - \alpha_0)$  についても同様)．すなわち, 本定理の主張は  $\mathbb{J}_{1,n}(\cdot), \mathbb{J}_{2,n}(\cdot)$  の確率評価に帰着する．その評価のため, 統計的確率場の多項式型大偏差不等式 (Yoshida, 2011) の理論を援用する．証明は複数のモーメント評価の検証を伴うが, 今モーメントの存在は必要なだけ仮定している状況であり (仮定 2.1, 3.1, 3.2), したがって一般性を失うことなく  $d = p_\gamma = p_\alpha = 1$  としてよいので, 以下その場合で見えていく．ここでは疑似尤度  $\mathbb{M}_n$  が混合収束率を持つことにより対応する統計的確率場の直接評価はできないため, 疑似スコア関数から構成される確率場を対象とした多項式型大偏差不等式を適用する．

仮定の下,  $\mathbb{G}_{1,n}(\cdot)$  に関しては, および正定数  $M, \epsilon$  が存在して

$$(5.1) \quad \begin{aligned} \max_{k \in \{1, 2, 3\}} \sup_{n \in \mathbb{N}} \left\{ E[|\sqrt{T_n} \mathbb{G}_{1,n}(\gamma_0)|^k] + E \left[ \sup_{\gamma \in \Theta_\gamma} |\partial_\gamma^k \mathbb{G}_{1,n}(\gamma)|^k \right] \right\} < \infty \quad (K > 0), \\ \sup_{n \in \mathbb{N}} \left\{ E \left[ \sup_{\gamma \in \Theta_\gamma} |\sqrt{T_n}(\mathbb{G}_{1,n}(\gamma) - \mathbb{G}_\gamma^\infty(\gamma))|^{M+\epsilon} \right] + E[|\sqrt{T_n}(\partial_\gamma \mathbb{G}_{1,n}(\gamma_0) - \mathcal{I}_\gamma)|^M] \right\} < \infty \end{aligned}$$

が成り立つことを示せば十分である．任意の  $k \in \{1, \dots, 4\}$  に対して

$$\mathbb{G}_{1,n}(\gamma_0) = \frac{1}{T_n} \sum_{j=1}^n \partial_\gamma S_{j-1}^{-1} \{(\Delta_j X)^2 - h_n S_{j-1}\}$$

$$\partial_\gamma^k \mathbb{G}_{1,n}(\gamma) = \frac{2}{n} \sum_{j=1}^n \partial_\gamma^k \left( \frac{\partial_\gamma c_{j-1}(\gamma)}{c_{j-1}(\gamma)} \right) + \frac{1}{T_n} \sum_{j=1}^n \partial_\gamma^{k+1} S_{j-1}^{-1}(\gamma) (\Delta_j X)^2$$

であることに注意すれば、(5.1)は Sobolev の不等式と Masuda and Uehara (2017, Lemma 5.3) の証明内の評価からしたがう。また、

$$\mathbb{G}_{1,n}(\gamma) = \frac{2}{n} \sum_{j=1}^n \frac{\partial_\gamma c_{j-1}(\gamma)}{c_{j-1}^3(\gamma)} (c_{j-1}^2(\gamma) - c_{j-1}^2) - \frac{2}{T_n} \sum_{j=1}^n \frac{\partial_\gamma c_{j-1}(\gamma)}{c_{j-1}^3(\gamma)} \{(\Delta_j X)^2 - h_n c_{j-1}^2\}$$

であり、Sobolev の不等式と Masuda (2013, Lemma 4.3), Masuda and Uehara (2017, Lemma 5.3) の証明内の評価より、任意の  $K > 0$  について

$$\sup_{n \in \mathbb{N}} E \left[ \sup_{\gamma \in \Theta_\gamma} |\sqrt{T_n}(\mathbb{G}_{1,n}(\gamma) - \mathbb{G}_\gamma^\infty(\gamma))|^K \right] < \infty$$

を得る。同様にして、

$$\partial_\gamma \mathbb{G}_{1,n}(\gamma_0) = \frac{4}{n} \sum_{j=1}^n \frac{(\partial_\gamma c_{j-1})^2}{c_{j-1}^2} - \frac{2}{T_n} \sum_{j=1}^n \left[ \frac{(\partial_\gamma^2 c_{j-1}) c_{j-1} - 3(\partial_\gamma c_{j-1})^2}{c_{j-1}^4} \{(\Delta_j X)^2 - h_n c_{j-1}^2\} \right]$$

なので

$$\sup_{n \in \mathbb{N}} E[|\sqrt{T_n}(\partial_\gamma \mathbb{G}_{1,n}(\gamma_0) - \mathcal{I}_\gamma)|^K] < \infty$$

を得る。以上より、Masuda (2013, Theorem 3.5) を適用して  $\sqrt{T_n}(\hat{\gamma}_n - \gamma_0)$  の裾確率評価

$$\sup_n P(|\sqrt{T_n}(\hat{\gamma}_n - \gamma_0)| > r) \leq \frac{C_K}{r^K}$$

を得る。

Taylor の公式から  $k \in \{0, 1, 2, 3\}$  について

$$\begin{aligned} \sqrt{T_n} \partial_\alpha^k \mathbb{G}_{2,n}(\alpha, \hat{\gamma}_n) &= \sqrt{T_n} \partial_\alpha^k \mathbb{G}_{2,n}(\alpha, \gamma_0) \\ &\quad + \left( \int_0^1 \partial_\gamma \partial_\alpha^k \mathbb{G}_{2,n}(\alpha, \gamma_0 + u(\hat{\gamma}_n - \gamma_0)) du \right) [\sqrt{T_n}(\hat{\gamma}_n - \gamma_0)] \end{aligned}$$

となる。すでに任意の  $K > 0$  に対して  $\sqrt{T_n}(\hat{\gamma}_n - \gamma_0)$  の  $L^K$ -有界性を確保しているため、右辺第 2 項は  $L^K$ -有界である。ゆえに  $\mathbb{G}_{2,n}(\alpha, \gamma_0)$  については、任意の  $K > 0$  とある正定数  $M, \epsilon$  に対して

$$\begin{aligned} \max_{k \in \{1, 2, 3\}} \sup_{n \in \mathbb{N}} \left\{ E[|\sqrt{T_n} \mathbb{G}_{2,n}(\alpha, \gamma_0)|^K] + E \left[ \sup_{\alpha \in \Theta_\alpha} |\partial_\alpha^k \mathbb{G}_{2,n}(\alpha, \gamma_0)|^K \right] \right\} &< \infty, \\ \sup_{n \in \mathbb{N}} \left\{ E \left[ \sup_{\alpha \in \Theta_\alpha} |\sqrt{T_n}(\mathbb{G}_{2,n}(\alpha, \gamma_0) - \mathbb{G}_\alpha^\infty(\alpha))|^{M+\epsilon} \right] + E[|\sqrt{T_n}(\partial_\alpha \mathbb{G}_{2,n}(\alpha, \gamma_0) - \mathcal{I}_\alpha)|^M] \right\} &< \infty \end{aligned}$$

を示せば十分である。今の場合

$$\begin{aligned} \mathbb{G}_{2,n}(\alpha, \gamma_0) &= \frac{1}{T_n} \sum_{j=1}^n S_{j-1}^{-1} \partial_\alpha a_{j-1}(\alpha, \gamma_0) (\Delta_j X - h_n a_{j-1}) \\ &\quad + \frac{1}{n} \sum_{j=1}^n S_{j-1}^{-1} \partial_\alpha a_{j-1}(\alpha, \gamma_0) (a_{j-1} - a_{j-1}(\alpha, \gamma_0)), \end{aligned}$$

$$\begin{aligned} \partial_\alpha^k \mathbb{G}_{2,n}(\alpha, \gamma_0) &= \frac{1}{T_n} \sum_{j=1}^n S_{j-1}^{-1} \partial_\alpha^{k+1} a_{j-1}(\alpha, \gamma_0) (\Delta_j X - h_n a_{j-1}(\alpha, \gamma_0)) \\ &\quad - \frac{1}{n} \sum_{j=1}^n S_{j-1}^{-1} \partial_\alpha^{k-1} (\partial_\alpha a_{j-1}(\alpha, \gamma_0))^2 \end{aligned}$$

であることに注意すれば,  $\sqrt{T_n}(\hat{\gamma}_n - \gamma_0)$  のときと同様にして裾確率評価

$$\sup_n P(|\sqrt{T_n}(\hat{\alpha}_n - \alpha_0)| > r) \leq \frac{C_K}{r^K}$$

を得る. 以上で (3.5) が示された.

## 5.2 漸近標準正規性 (3.6) の証明

裾確率評価 (3.5) から  $\sqrt{T_n}(\hat{\theta}_n - \theta_0) = O_p(1)$ , 特に一貫性  $\hat{\theta}_n \xrightarrow{P} \theta_0$  がしたがうことに注意する. まず

$$(5.2) \quad \mathbb{G}_n(\theta_0) := \left( \mathbb{G}_{2,n}(\alpha_0, \hat{\gamma}_n) + \left( \frac{1}{n} \sum_{j=1}^n M'_{j-1}[\partial_\gamma a_{j-1}] \right) [\hat{\gamma}_n - \gamma_0], \mathbb{G}_{1,n}(\gamma_0) \right)$$

について

$$(5.3) \quad \sqrt{T_n} \mathbb{G}_n(\theta_0) \xrightarrow{\mathcal{L}} N(0, \Sigma)$$

を示す.  $\sqrt{T_n} \mathbb{G}_n(\theta_0)$  の第一成分について, Taylor 展開から

$$\begin{aligned} \sqrt{T_n} \mathbb{G}_{2,n}(\alpha_0, \hat{\gamma}_n) &= \sqrt{T_n} \mathbb{G}_{2,n}(\theta_0) + \partial_\gamma \mathbb{G}_{2,n}(\theta_0) [\sqrt{T_n}(\hat{\gamma}_n - \gamma_0)] \\ &\quad + \frac{1}{\sqrt{T_n}} \left( \int_0^1 \partial_\gamma^{\otimes 2} \mathbb{G}_{2,n}(\alpha_0, \gamma_0 + u(\hat{\gamma}_n - \gamma_0)) du \right) [(\sqrt{T_n}(\hat{\gamma}_n - \gamma_0))^{\otimes 2}] \end{aligned}$$

を得る. Sobolev の不等式より  $\int_0^1 \partial_\gamma^{\otimes 2} \mathbb{G}_{2,n}(\alpha_0, \gamma_0 + u(\hat{\gamma}_n - \gamma_0)) du = O_p(1)$  が示され, 右辺第三項は  $o_p(1)$  となる. また, モーメント条件の下,  $a$  の Lipschitz 連続性および Masuda (2013, Lemma 4.5) から  $E^{j-1}[\|\int_j (a_s - a_{j-1}) ds\|^2] \lesssim h_n^3$ , さらに  $\int_j c_{s-} dJ_s$  のマルチンゲール性により,  $E^{j-1}[\int_j c_{s-} dJ_s] = 0$ ,  $E^{j-1}[\|\int_j c_{s-} dJ_s\|^2] \lesssim h_n$  が従うため, Genon-Catalot and Jacod (1993, Lemma 9) より

$$\begin{aligned} \partial_\gamma \mathbb{G}_{2,n}(\theta_0) &= \frac{1}{T_n} \sum_{j=1}^n \partial_\gamma M'_{j-1} [\Delta_j X - h_n a_{j-1}] - \frac{1}{n} \sum_{j=1}^n M'_{j-1} [\partial_\gamma a_{j-1}] \\ &= o_p(1) - \frac{1}{n} \sum_{j=1}^n M'_{j-1} [\partial_\gamma a_{j-1}] \end{aligned}$$

が成り立つ. これを踏まえて

$$\begin{aligned} \sqrt{T_n} \mathbb{G}_n(\theta_0) &= (\sqrt{T_n} \mathbb{G}_{2,n}(\theta_0), \sqrt{T_n} \mathbb{G}_{1,n}(\gamma_0)) + o_p(1) \\ &= \left( \frac{1}{\sqrt{T_n}} \sum_{j=1}^n M'_{j-1} [c_{j-1} \Delta_j J], \right. \\ &\quad \left. \frac{1}{\sqrt{T_n}} \sum_{j=1}^n \{h_n \text{trace}(S_{j-1}^{-1} \partial_\gamma S_{j-1}) - M''_{j-1} [(c_{j-1} \Delta_j J)^{\otimes 2}]\} \right) + o_p(1) \end{aligned}$$

を得る. 最右辺第一項は各  $n$  について  $(\mathcal{F}_{t_j})$ -マルチンゲール差分になっていることに注意する.

モーメント評価  $E[|J_{h_n}|^q] \lesssim h_n$  により, 最右辺第一項について Lyapunov 条件が容易に示される. 以下  $u \in \mathbb{R}^{p\alpha}$ ,  $v \in \mathbb{R}^{p\gamma}$  を任意に固定し, 二次特性量を計算する.

- $E[(\Delta_j J)^{\otimes 2}] = h_n I_r$  より

$$\begin{aligned} & \left( \frac{1}{T_n} \sum_{j=1}^n E^{j-1} [(M'_{j-1} \otimes M'_{j-1}) [c_{j-1} \Delta_j J, c_{j-1} \Delta_j J]] \right) [u, u'] \\ &= \frac{1}{n} \sum_{j=1}^n M'_{j-1} [u, c_{j-1}] \times M'_{j-1} [u', c_{j-1}] \\ & \xrightarrow{P} \Sigma_\alpha [u, u']. \end{aligned}$$

- 第二要素について, まず

$$\begin{aligned} & \left( \frac{1}{T_n} \sum_{j=1}^n E^{j-1} [\{M''_{j-1} [h_n S_{j-1} - (c_{j-1} \Delta_j J)^{\otimes 2}]\}^{\otimes 2}] \right) [v, v'] \\ &= - \left( \frac{h_n}{n} \sum_{j=1}^n (M''_{j-1} \otimes M''_{j-1}) [S_{j-1}, S_{j-1}] \right) [v, v'] \\ & \quad + \left( \frac{1}{T_n} \sum_{j=1}^n E^{j-1} [(M''_{j-1} \otimes M''_{j-1}) [(c_{j-1} \Delta_j J)^{\otimes 2}, (c_{j-1} \Delta_j J)^{\otimes 2}] \right) [v, v'] \\ &= \left( \frac{1}{T_n} \sum_{j=1}^n E^{j-1} [(M''_{j-1} \otimes M''_{j-1}) [(c_{j-1} \Delta_j J)^{\otimes 2}, (c_{j-1} \Delta_j J)^{\otimes 2}] \right) [v, v'] + o_p(1) \end{aligned}$$

に注意する. Lévy-Khintchin 表現

$$\log E[e^{iu^T J_t}] = t \int_{\mathbb{R}^r} (e^{iu^T z} - 1 - iu^T z) \nu_0(dz)$$

を偏微分して

$$\begin{aligned} E[J_t^{(i_1)} J_t^{(i_2)} J_t^{(i_3)}] &= \nu_{i_1, i_2, i_3}(3)t \\ E[J_t^{(i_1)} J_t^{(i_2)} J_t^{(i_3)} J_t^{(i_4)}] &= \nu_{i_1, i_2, i_3, i_4}(4)t + O(t^2) \end{aligned}$$

を得る. これを適用することで,

$$\begin{aligned} & \left( \frac{1}{T_n} \sum_{j=1}^n E^{j-1} [(M''_{j-1} \otimes M''_{j-1}) [(c_{j-1} \Delta_j J)^{\otimes 2}, (c_{j-1} \Delta_j J)^{\otimes 2}] \right) [v, v'] \\ &= \frac{1}{T_n} \sum_{j=1}^n E^{j-1} [M''_{j-1} [v, (c_{j-1} \Delta_j J)^{\otimes 2}] \times M''_{j-1} [v', (c_{j-1} \Delta_j J)^{\otimes 2}] \\ &= \frac{1}{T_n} \sum_{j=1}^n E^{j-1} \left[ \sum_{i_1, i_2=1}^d (M''_{j-1} [v])^{(i_1 i_2)} \left( \sum_{s=1}^r c_{j-1}^{(i_1 s)} (\Delta_j J)^{(s)} \right) \left( \sum_{t=1}^r c_{j-1}^{(i_2 t)} (\Delta_j J)^{(t)} \right) \right. \\ & \quad \left. \times \sum_{i'_1, i'_2=1}^d (M''_{j-1} [v'])^{(i'_1 i'_2)} \left( \sum_{s'=1}^r c_{j-1}^{(i'_1 s')} (\Delta_j J)^{(s')} \right) \left( \sum_{t'=1}^r c_{j-1}^{(i'_2 t')} (\Delta_j J)^{(t')} \right) \right] \\ &= \frac{1}{T_n} \sum_{j=1}^n \left\{ \sum_{s, t, s', t'=1}^r (E^{j-1} [(\Delta_j J)^{(s)} (\Delta_j J)^{(t)} (\Delta_j J)^{(s')} (\Delta_j J)^{(t')}] \right\} \end{aligned}$$

$$\begin{aligned}
& \times \left( \sum_{i_1, i_2, i'_1, i'_2=1}^d (M''_{j-1}[v])^{(i_1 i_2)} c_{j-1}^{(i_1 s)} c_{j-1}^{(i_2 t)} \times (M''_{j-1}[v'])^{(i'_1 i'_2)} c_{j-1}^{(i'_1 s')} c_{j-1}^{(i'_2 t')} \right) \Big\} \\
& = \frac{1}{n} \sum_{j=1}^n \left( \sum_{s, t, s', t'=1}^r \nu_{s, t, s', t'}(4) \times M''_{j-1}[v, c_{j-1}^{(\cdot s)}, c_{j-1}^{(\cdot t)}] \times M''_{j-1}[v', c_{j-1}^{(\cdot s')}, c_{j-1}^{(\cdot t')}] \right) \\
& \xrightarrow{P} \Sigma_\gamma[v, v'].
\end{aligned}$$

•  $E[|J_{h_n}|^2] = O_p(h_n)$  より

$$\frac{1}{n} \sum_{j=1}^n (M'_{j-1}[c_{j-1} \Delta_j J] \otimes \text{trace}(S_{j-1}^{-1} \partial_\gamma S_{j-1}))[u, v] = o_p(1)$$

であり, さらに

$$\begin{aligned}
& - \frac{1}{T_n} \sum_{j=1}^n E^{j-1} [(M'_{j-1}[c_{j-1} \Delta_j J] \otimes M''_{j-1}[(c_{j-1} \Delta_j J)^{\otimes 2}]) [u, v]] \\
& = - \frac{1}{T_n} \sum_{j=1}^n E^{j-1} \left[ \left( \sum_{s=1}^d (M'_{j-1})^{(\cdot s)} (c_{j-1} \Delta_j J)^{(s)} \right) \right. \\
& \quad \left. \otimes \left( \sum_{t, u=1}^d (M''_{j-1})^{(\cdot tu)} ((c_{j-1} \Delta_j J)^{\otimes 2})^{(tu)} \right) [u, v] \right] \\
& = - \frac{1}{T_n} \sum_{j=1}^n E^{j-1} \left[ \sum_{s, t, u=1}^d \sum_{k_1, k_2, k_3}^r \Delta_j J^{(k_1)} \Delta_j J^{(k_2)} \Delta_j J^{(k_3)} \right. \\
& \quad \left. \times c_{j-1}^{(sk_1)} c_{j-1}^{(tk_2)} c_{j-1}^{(uk_3)} ((M'_{j-1})^{(\cdot s)} \otimes (M''_{j-1})^{(\cdot tu)}) [u, v] \right] \\
& = - \frac{1}{n} \sum_{j=1}^n \sum_{k_1, k_2, k_3}^r \nu_{k_1, k_2, k_3}(3) \sum_{s, t, u=1}^d c_{j-1}^{(sk_1)} c_{j-1}^{(tk_2)} c_{j-1}^{(uk_3)} ((M'_{j-1})^{(\cdot s)} \otimes (M''_{j-1})^{(\cdot tu)}) [u, v] \\
& \xrightarrow{P} \Sigma_{\alpha, \gamma}[u, v].
\end{aligned}$$

以上の下, マルチンゲール中心極限定理 (e.g. Dvoretzky, 1972) から (5.3) がしたがう.

これまでの評価を踏まえて, 行列

$$\begin{aligned}
\mathcal{I}_\alpha(\theta) &:= \int (\partial_\alpha M'(x, \theta)[a(x, \theta_0) - a(x, \theta)] - S^{-1}(x, \gamma)[(\partial_\alpha a(x, \theta))^{\otimes 2}]) \pi_0(dx), \\
\mathcal{I}_\gamma(\gamma) &:= \int \text{trace} (2(S^{-1} \partial_\gamma S S^{-1} \partial_\gamma S S^{-1})(x, \gamma) S(x, \gamma_0) - (S^{-1} \partial_\gamma S S^{-1} \partial_\gamma S)(x, \gamma) \\
& \quad + (S^{-1} \partial_\gamma^{\otimes 2} S)(x, \gamma) - (S^{-1} \partial_\gamma^{\otimes 2} S S^{-1})(x, \gamma) S(x, \gamma_0)) \pi_0(dx), \\
B(\theta) &:= - \int \{ \partial_\gamma M'(x, \theta)[a(x, \theta_0) - a(x, \theta)] + S^{-1}(x, \gamma_0)[\partial_\alpha a(x, \theta), \partial_\gamma a(x, \theta)] \} \pi_0(dx)
\end{aligned}$$

に対して一様収束

$$\begin{aligned}
\sup_{\gamma \in \Theta_\gamma} |\partial_\gamma \mathbb{G}_{1, n}(\gamma) - \mathcal{I}_\gamma(\gamma)| &\xrightarrow{P} 0, \\
\sup_{\theta \in \Theta} |\partial_\alpha \mathbb{G}_{2, n}(\theta) - \mathcal{I}_\alpha(\theta)| &\xrightarrow{P} 0,
\end{aligned}$$

$$\sup_{\theta \in \Theta} |\partial_\gamma \mathbb{G}_{2,n}(\theta) - B(\theta)| \xrightarrow{P} 0$$

が成り立つことがわかる。これと  $\mathbb{G}_n$  の定義(5.2), (5.3) より

$$(5.4) \quad \sqrt{T_n} \Sigma^{-1/2} \begin{pmatrix} \partial_\alpha \mathbb{G}_{2,n}(\hat{\theta}_n) & \partial_\gamma \mathbb{G}_{2,n}(\hat{\theta}_n) \\ 0 & \partial_\gamma \mathbb{G}_{1,n}(\hat{\gamma}_n) \end{pmatrix} \begin{pmatrix} \hat{\alpha}_n - \alpha_0 \\ \hat{\gamma}_n - \gamma_0 \end{pmatrix} \xrightarrow{\mathcal{L}} N(0, I_p)$$

を得る。例えば  $\hat{\Sigma}_{\alpha,n}$  について Taylor 展開を適用すると

$$\hat{\Sigma}_{\alpha,n}[u, u'] = \frac{1}{n} \sum_{j=1}^n S_{j-1}^{-1} [\partial_\alpha a_{j-1}[u], \partial_\alpha a_{j-1}[u']] + o_p \left( \frac{1}{\sqrt{T_n}} \right)$$

が得られる。これにエルゴード定理と Genon-Catalot and Jacod (1993, Lemma 9) を適用すれば  $\hat{\Sigma}_{\alpha,n} \xrightarrow{P} \Sigma_\alpha$  が従う。同様にして  $\hat{\Sigma}_{\gamma,n} \xrightarrow{P} \Sigma_\gamma$  および  $\hat{\Sigma}_{\alpha,\gamma,n} \xrightarrow{P} \Sigma_{\alpha,\gamma}$  も示され, (3.5), (5.4) と Slutsky の補題より (3.6) を得る。

### 5.3 モーメント収束(3.7)の証明

すでに示した  $\hat{\Sigma}_n \xrightarrow{P} \Sigma$  および  $\hat{\mathcal{I}}_n \xrightarrow{P} \mathcal{I}$  と (3.6) により,

$$\sqrt{T_n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N(0, \mathcal{I}^{-1} \Sigma (\mathcal{I}^{-1})^\top)$$

が成り立つ。(3.5) から任意の  $q > 0$  に対する  $\{|\sqrt{T_n}(\hat{\theta}_n - \theta_0)|^q\}_{n \in \mathbb{N}}$  の一様可積分性が成り立つので(3.7)が得られる。

## 謝 辞

丁寧に査読していただき、多くの改善点の指摘および有益なコメントを下された査読者に御礼申し上げます。本研究は JST CREST, JSPS 科研費 JP26400204(増田)の助成を受けたものです。

## 参 考 文 献

- Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type, *Finance and Stochastics*, **2**(1), 41–68.
- Brouste, A., Fukasawa, M., Hino, H., Iacus, S., Kamatani, K., Koike, Y., Masuda, H., Nomura, R., Ogihara, T., Shimizu, Y., Uchida, M. and Yoshida, N. (2014). The YUIMA project: A computational framework for simulation and inference of stochastic differential equations, *Journal of Statistical Software*, **57**(4), 1–51.
- Dvoretzky, A. (1972). Asymptotic normality for sums of dependent random variables, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (University of California, Berkeley, California, 1970/1971)*, Vol. II: Probability Theory, 513–535, University of California Press, Berkeley, California.
- Genon-Catalot, V. and Jacod, J. (1993). On the estimation of the diffusion coefficient for multi-dimensional diffusion processes, *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*, **29**, 119–151.
- Kamatani, K. and Uchida, M. (2015). Hybrid multi-step estimators for stochastic differential equations based on sampled data, *Statistical Inference for Stochastic Processes*, **18**(2), 177–204.
- Masuda, H. (2013). Convergence of Gaussian quasi-likelihood random fields for ergodic Lévy driven



- SDE observed at high frequency, *Annals of Statistics*, **41**, 1593–1641.
- Masuda, H. and Uehara, Y. (2017). Two-step estimation of ergodic Lévy driven SDE, *Statistical Inference for Stochastic Processes*, **20**(1), 105–137.
- Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, U.K.
- Uchida, M. and Yoshida, N. (2012). Adaptive estimation of an ergodic diffusion process based on sampled data, *Stochastic Processes and Their Applications*, **122**(8), 2885–2924.
- Yoshida, N. (2011). Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations, *Annals of the Institute of Statistical Mathematics*, **63**(3), 431–479.

## On Stepwise Estimation of Lévy Driven Stochastic Differential Equation

Yuma Uehara and Hiroki Masuda

Graduate School of Mathematics, Kyushu University

We consider estimation of a non-Gaussian Lévy driven stochastic differential equation. Under high-frequency data and exponential ergodicity, we propose the stepwise estimation procedure based on a Gaussian quasi-score function: first we estimate the scale parameter while ignoring the drift coefficient, and then focus on the drift parameter by plugging in the estimated scale parameters, and derive the asymptotic normality and the tail probability estimate of the proposed estimators. This stepwise strategy not only reduces computational cost but may also stabilize estimate accuracy. Unlike the diffusion case, the asymptotic covariance matrix associated with the drift parameter takes a different form when there is a common parameter in the coefficients.