

基于离散观测的 OU 型过程的经验似然估计

孙曙光, 张新生

复旦大学管理学院统计学系, 上海 200433

E-mail: shuguang@gmail.com, xszhang@fudan.edu.cn

收稿日期: 2007-08-20; 接受日期: 2008-07-03

国家自然科学基金 (批准号: 10671037) 和上海市教委科研项目 (批准号: 06FZ035) 资助项目

摘要 基于离散观测样本, 借助条件特征函数, 提出了 OU 型过程参数的经验似然估计, 并证明了最大经验似然估计的相合性和渐近正态性, 同时证明了在适当的附加条件下, 该估计是渐近有效的. 当驱动 Lévy 过程具有某种特殊形式时, 发现 OU 型过程的强度参数能够首先估计出来. 在此特殊情形下, 提出了 OU 型过程中其余参数的最大经验似然估计, 并讨论了其相合性、渐近正态性和渐近有效性. 基于经验似然比统计量, 给出了参数检验和估计方程检验的似然比检验方法. 模拟显示所提出的估计方法是准确和稳定的.

关键词 OU 型过程 条件特征函数 经验似然 工具变量

MSC(2000) 主题分类 62M02, 62M05

1 引言

称随机过程 $X = \{X(t), t \geq 0\}$ 是 Ornstein-Uhlenbeck (OU) 型的, 如果它满足随机微分方程

$$\begin{cases} dX(t) = -cX(t)dt + dZ(ct), \\ X(0) = X_0, \end{cases} \quad (1.1)$$

其中 $c > 0$ 是强度参数, $Z = \{Z(t), t \geq 0\}$ 是背景驱动 Lévy 过程, 并由参数向量 $\tilde{\theta}$ 唯一确定, 初始条件 X_0 与 Z 无关.

OU 型过程在金融和计量经济等领域有着广泛应用, 例如 Barndorff-Nielsen 和 Shephard^[1] 提出了利用 Gamma-OU 和 IG-OU 等 OU 型过程刻画金融资产的波动情形以反映突发事件的影响.

但在实际应用中, 连续时间的随机过程只能在离散的时间上观测. 在过去 40 年间基于离散观测的统计方法得到广泛发展. 最大似然估计是最常用的一种方法. 除少数特殊过程外, OU 型过程的似然函数没有简单的闭形式. 通常, 似然函数的计算或通过计算 Fokker-Planck-Kolmogorov 偏微分方程的数值解 (参见文献 [2]), 或通过对轨道的大量精细模拟抽样来实现 (参见文献 [3, 4]). 对于 IG-OU 过程或 NIG-OU 过程, Barndorff-Nielsen^[5] 提出了模

拟似然的方法; 对于 Gamma-OU 过程, 张等^[6] 通过特征函数的 Fourier 逆变换得到转移密度函数, 进一步通过 Gaver-Stefest 算法构造了似然函数的逼近序列, 从而得到最大似然估计; Valdivieso 等^[7] 对一般的 OU 型过程提出了基于特征函数的传统或分数快速 Fourier 变换 (FFT) 的最大似然方法, 但是没有给出统计量理论性质. Jongbloed 等^[8] 采用非参数方法对 OU 型过程的估计和检验问题进行了研究.

相对于转移密度函数, OU 型过程的条件特征函数更容易得到, 因此基于特征函数的统计方法是最大似然方法之外的重要选择. 针对独立同分布随机变量序列及 ARMA 序列, Feuerverger 和 McDunnough^[9] 以及 Feuerverger^[10] 建立了基于特征函数的估计方法. 而对连续时间模型, Singleton^[11], Jiang 和 Knight^[12], 以及 Chacko 和 Viceira^[13] 通过条件特征函数构造矩条件, 研究了 GMM 估计的性质.

经验似然方法是近 20 年来统计领域最重要的成就之一. Owen^[14, 15] 通过经验似然方法构造了非参数置信区间, 之后 Qin 和 Lawless^[16] 将经验似然方法应用于估计方程问题. 若样本是独立的, 经验似然方法与最大似然方法有很多相似的性质, 例如对数经验似然比统计量也是渐近卡方分布的. 经验似然方法在半参数模型领域也有重要应用, 例如文献 [17, 18]. Owen^[19] 对经验似然方法的性质和应用有更进一步的介绍. 对于独立同分布随机变量序列和 Lévy 过程, Kunitomo 和 Owada^[20] 通过特征函数构造矩条件, 证明了如果矩条件数随着样本量增大而增加, 可以同时估计方程和稳定分布的参数进行估计和检验.

本文的模型 OU 型过程是弱相关的, 并且当强度参数未知时, 无法通过离散观测构造一个独立同分布随机变量序列. 基于条件特征函数的鞅估计方程在这种情形下很有用. 本文假设有离散观测 $X_k = X(t_k)$, $\{t_k = k\Delta, k = 0, 1, \dots, n\}$, 观测时间间隔 $\Delta > 0$ 固定. 我们给出了 $\theta = (c, \tilde{\theta})$ 的基于条件特征函数的最大经验似然估计, 同时讨论了该估计的渐近性质, 如相合性、渐近正态性及渐近有效性. 工具变量方法也是必需的, 本文重点考察单指标 (SI) 和双指标 (DI) 两种工具变量, 并给出了能够得到渐近有效估计的最优工具变量. 当背景驱动 Lévy 过程 Z 为 A 型或 B 型时 (参见文献 [21, 定义 11.19]), OU 型过程的转移密度函数关于强度参数不可导, 不过我们证明了只要样本量足够大, 就有可能准确的找出强度参数 c 的真值, 在此基础上可以构造新的最大似然估计, 并且我们研究了该最大经验似然估计的性质, 并证明在适当的附加条件下, 该估计是渐近有效的. 针对参数的检验问题 ($H_0 : \theta = \theta_0$) 和估计方程的检验问题 ($H_0 : E_{\theta_0}[g(X_1, X_0, \theta_0)] = 0$), 我们分别建立了基于经验似然比统计量的检验方法.

2 OU 型过程

由随机微分方程 (1.1) 决定的 OU 型过程中, 强度参数 $c > 0$, Z 是一元 Lévy 过程. 文献 [21] 证明了 Z 存在右连左极的修正, 本文仅考虑这一修正过程. 根据左连续可料过程对 Lévy 过程的积分理论, 易证 OU 型方程 (1.1) 存在强解

$$X(t) = e^{-ct} X(0) + \int_0^t e^{-c(t-s)} dZ(cs),$$

对于任一 $\Delta > 0$, 可以递归的表示为

$$X(t + \Delta) = e^{-c\Delta} X(t) + \int_t^{t+\Delta} e^{-c(t+\Delta-s)} dZ(cs). \quad (2.1)$$

且 \mathbf{X} 是强 Markov 过程, 其转移密度函数可由条件特征函数确定.

这里有两个 OU 型过程的例子.

例 2.1 (Gamma-OU 过程) Gamma-OU 过程的平稳分布是以 $a > 0, b > 0$ 为参数的 Gamma 分布, 密度函数为

$$f_X(x) = \frac{b^a x^{a-1}}{\Gamma(a)} \exp(-bx), \quad \forall x > 0,$$

背景驱动 Lévy 过程 \mathbf{Z} 有累积函数 $\psi_{Z(1)}(t) = a\left(\frac{b}{b-it} - 1\right)$.

例 2.2 (IG-OU 过程) IG-OU 过程的平稳分布是以 $a > 0, b > 0$ 为参数的逆 Gauss (IG) 分布, 密度函数为

$$f_X(x) = \frac{a \exp(ab)}{\sqrt{2\pi x^3}} \exp\left(-\frac{1}{2}\left(\frac{a^2}{x} + b^2 x\right)\right), \quad \forall x > 0.$$

背景驱动 Lévy 过程 \mathbf{Z} 有累积函数

$$\psi_{Z(1)}(t) = \frac{ait}{\sqrt{b^2 - 2it}}.$$

以下是 OU 型过程的一些基本性质:

引理 2.1 OU 型方程 (1.1) 存在时齐的转移函数 $P_t(x, B)$, 满足

$$\int_{-\infty}^{\infty} e^{izy} P_t(x, dy) = \exp\left\{ize^{-ct}x + c \int_0^t \psi(e^{-cs}z) ds\right\},$$

其中 $\psi(z) = \log(\phi_{Z(1)}(z))$ 是 $Z(1)$ 的累积函数. 对任一 $\Delta > 0$, 记 $Y(t, \Delta) = \int_t^{t+\Delta} e^{-c(t+\Delta-s)} dZ(cs)$, 则

$$Y(t, \Delta) \stackrel{d}{=} Y(0, \Delta) = \int_0^\Delta e^{-c(\Delta-s)} dZ(cs).$$

文献 [21] 的定理 17.5 给出了 OU 型过程平稳的充分必要条件.

命题 2.1 对任一 $c > 0$ 固定, 如果 OU 型过程 \mathbf{X} 的生成元 (σ, γ, ν, c) 中 Lévy 测度 ν 满足

$$\int_{|x|>2} \log(|x|) d\nu(x) < \infty, \quad (2.2)$$

则 \mathbf{X} 存在唯一的平稳分布 μ , 且 μ 是自可分解分布.

令 $(\sigma_0, \gamma_0, \nu_0)$ 是平稳分布 μ 的 Lévy 三元组, 则

$$\sigma_0 = \frac{\sigma}{2}, \quad \gamma_0 = \gamma + \int_{|x|>1} \frac{x}{|x|} \omega(x) dx,$$

其中 ω 是 ν 的 Lévy 密度. 如果 ν_0 的 Lévy 密度 u 可微, 则 $\omega(x) = -u(x) - xu'(x)$ (参见文献 [5]).

对最大经验似然估计的渐近性质研究主要依赖于 OU 型过程的 α 混合的性质. 关于 α 和 β 混合的定义见文献 [22, 23]. 因为 $2\alpha \leq \beta \leq 1$, 所以 β 混合意味着 α 混合. 显然, 如果条件 (2.2) 成立, 则 OU 型过程 \mathbf{X} 是 β 混合的. 在稍强的条件下, OU 型过程是指数 β 混合的 (参见文献 [24]).

命题 2.2 设 OU 型过程 \mathbf{X} 的初始条件 X_0 与背景驱动 Lévy 过程 \mathbf{Z} 独立, \mathbf{Z} 的 Lévy 测度是 σ 有限的. 如果强度参数 $c > 0$, \mathbf{X} 是严平稳的, 并且存在 $q > 0$, 自可分解极限分布 q 阶绝对矩存在有限, 则 \mathbf{X} 是指数 β 混合且遍历的.

对 α 混合的随机变量, 有两个重要的矩不等式: Davydov 不等式和 Billingsley 不等式 (参见文献 [22, 25]). α 混合的随机变量序列还成立 Lindeberg 型中心极限定理 (参见文献 [23]).

3 OU 过程参数的最大经验似然估计

3.1 基于条件特征函数的估计方程

设条件 (2.2) 成立, 即 OU 型过程 \mathbf{X} 是严平稳的. 在离散时间 $\{t_k = k\Delta, k = 0, 1, 2, \dots, n\}$ 进行观测, 时间间隔 $\Delta > 0$ 固定. 简记 $X_k = X(k\Delta)$, 设 (x_0, x_1, \dots, x_n) 是 (X_0, X_1, \dots, X_n) 样本观测值. 记 p 元未知参数为 $\theta = (c, \tilde{\theta})$.

引理 3.1 记 $\{Y_k = X_k - e^{-c\Delta}X_{k-1}, k \in \mathbb{N}\}$, 则 $\{Y_1, Y_2, \dots\}$ 是独立同分布随机变量序列, 且与 $\int_0^\Delta e^{-c(\Delta-s)}dZ(cs)$ 同分布.

由引理 2.1 可得条件特征函数的闭形式. 记 $\int_0^\Delta e^{-c(\Delta-s)}dZ(cs)$ 的特征函数为 $\phi_\theta(z)$, 即

$$\phi_\theta(z) = \exp \left\{ c \int_0^t \psi_{\tilde{\theta}}(e^{-c\Delta s}z) ds \right\},$$

其中 $\psi_{\tilde{\theta}}(z) = \log(\phi_{Z(1)}(z))$ 是 $Z(1)$ 的累积函数, 则 OU 型过程 \mathbf{X} 的条件特征函数可以表示为

$$\phi_\theta(z|x) = e^{ize^{-c\Delta}x} \phi_\theta(z).$$

据此可以建立估计方程

$$h(X_k, X_{k-1}, \theta) = e^{izX_k} - \phi_\theta(z|X_{k-1}) = e^{ize^{-c\Delta}X_{k-1}} [e^{iz(X_k - e^{-c\Delta}X_{k-1})} - \phi_\theta(z)].$$

在随机微分方程 (1.1) 中, $Z(ct)$ 的特殊时速 ct 使得 \mathbf{X} 的平稳分布不依赖于强度参数 c . 在一般情形下, $\lim_{c \rightarrow \infty} \phi_\theta(z) = \phi_{X, \tilde{\theta}}(z)$, $\lim_{c \rightarrow \infty} \frac{\partial \phi_\theta(z)}{\partial c} = 0$, 其中 $\phi_{X, \tilde{\theta}}(z)$ 是平稳分布的特征函数, 因此方程 $E_\theta h(X_k, X_{k-1}, \theta) = 0$ 的解并不唯一, 强度参数 c 很难估计.

在这种情形下, 利用工具变量是一种有效的解决方法. 令 $\gamma_\theta(\tau, x)$ 是有界的工具函数, 满足基本条件: 数学期望 $\mathbb{E}_{\theta_0}[\gamma_{\theta_0}(\tau, X_0)e^{i\tau e^{-c_0\Delta}X_0}] \neq 0$, 且 $\lim_{c \rightarrow \infty} \gamma_\theta(\tau, X_0)$ 关于 X_0 是非退化的. 令

$$\tilde{g}(\tau, X_k, X_{k-1}, \theta) = \gamma_\theta(\tau, X_{k-1}) [\exp(i\tau X_k) - \phi_\theta(\tau|X_{k-1})].$$

由引理 3.1 知, $\{\tilde{g}(\tau, X_k, X_{k-1}, \theta_0), k = 1, \dots, n\}$ 是鞅差序列, $\mathbb{E}_{\theta_0} \tilde{g}(\tau, X_1, X_0, \theta_0) = 0$.

有两个问题需要考虑: 离散化参数 τ 的选择, 工具函数 $\gamma_\theta(\tau, x)$ 的选择. 选择不同的 τ 的集合, 会得到不同的 $\hat{\theta}$ 的值. 当集合 $\{\tau\}$ 有限、固定时, Qin 和 Lawless^[16] 证明了在独立同分布样本下基于广义估计方程 (GEE) 的最大经验似然估计的渐近性质. 他们的充分条件可以在本文模型下得到验证. 当然本文的观测是弱相依的, 而 OU 型过程的 (指数) β 混合性可以保证最大经验似然估计的渐近性质.

当离散化参数 τ 的数目随着样本量增加而增大时, 选择合适的工具函数可以使最大经验似然估计达到 Cramér-Rao 下界. 本文主要考虑两类工具函数构造的估计方程:

1. 双指标矩函数: 取 $\tau = (s, z) \in \mathbb{R}^2$, $\gamma_\theta(s, x) = \overline{\gamma_\theta(-s, x)}$, 令

$$\tilde{g}(\tau, X_k, X_{k-1}, \theta) = \gamma_\theta(s, X_{k-1}) [e^{izX_k} - \phi_\theta(z|X_{k-1})].$$

2. 单指标矩函数: 取 $\tau = z \in \mathbb{R}$, $\gamma_\theta(z, x) = \overline{\gamma_\theta(-z, x)}$, 令

$$\tilde{g}(z, X_k, X_{k-1}, \theta) = \gamma_\theta(z, X_{k-1}) [e^{izX_k} - \phi_\theta(z|X_{k-1})],$$

本文将证明最优双指标工具函数为 $\gamma_\theta^{DI}(s, x) = e^{isx}$, 最优单指标工具函数为 $\gamma_\theta^{SI}(t, x) = \frac{1}{2\pi} \int e^{-ity} \frac{\partial}{\partial \theta} \ln p_\theta(y; x, \Delta) dy$. “最优”是指, 当采用这一工具函数时, 得到的最大经验似然估计是渐近有效的. 易见最优单指标工具函数与未知的转移密度函数有关, 因此本文主要讨论双指标工具函数.

现在详细说明如何采用双指标工具函数构造估计方程. 记条件特征函数 $\phi_\theta(z|x)$ 的实部与虚部分别为 $\phi_\theta^R(z|x)$ 和 $\phi_\theta^I(z|x)$, 工具函数 $\gamma_\theta(s, x)$ 的实部和虚部分别为 $\gamma_\theta^R(s, x)$ 和 $\gamma_\theta^I(s, x)$. 取 $m_1 \geq 2$, 在 m_1 个不同点上计算条件特征函数的值, 离散点为 $z = z_l, l = 1, 2, \dots, m_1$, 且 $z_1 < z_2 < \dots < z_{m_1}$. 在 m_2 个点上对工具函数进行离散化, 离散点为 $s = s_l, l = 1, 2, \dots, m_2$, 且 $s_1 < s_2 < \dots < s_{m_2}$. m_1 和 m_2 的选择非常重要, 可以依赖于样本量 n . 特别地, 我们取 $m_1 m_2 = m(n) = [n^{\frac{1}{2}-\eta}]$, 其中 $0 < \eta < \frac{1}{2}$, $[n^{\frac{1}{2}-\eta}]$ 是不大于 $n^{\frac{1}{2}-\eta}$ 的最大整数. 定义

$$g(y, x, \theta, s, z) = (\gamma_\theta^R(s, x)[\cos(zy) - \phi_\theta^R(z|x)], \gamma_\theta^R(s, x)[\sin(zy) - \phi_\theta^I(z|x)], \\ \gamma_\theta^I(s, x)[\cos(zy) - \phi_\theta^R(z|x)], \gamma_\theta^I(s, x)[\sin(zy) - \phi_\theta^I(z|x)]^T.$$

记 $4m_1 m_2$ 元向量

$$\mathbf{g}(X_k, X_{k-1}, \theta) = (g(X_k, X_{k-1}, \theta, s_1, z_1)^T, g(X_k, X_{k-1}, \theta, s_1, z_{m_1})^T, \dots, \\ g(X_k, X_{k-1}, \theta, s_{m_2}, z_1)^T, g(X_k, X_{k-1}, \theta, s_{m_2}, z_{m_1})^T)^T,$$

由引理 3.1, $\{\mathbf{g}(X_k, X_{k-1}, \theta_0), k = 1, \dots, n\}$ 是鞅差阵列, 数学期望 $\mathbb{E}_{\theta_0} \mathbf{g}(X_1, X_0, \theta_0) = 0$.

3.2 经验似然估计

Owen^[14, 15] 最早给出经验似然方法的渐近性质, Qin 和 Lawless^[16] 将之推广到估计方程. 为同时估计参数 c 和 $\tilde{\theta}$, 本文构造了类似于 Kunitomo 和 Owada^[20] 的经验似然估计. 由引理 3.1, (Y_1, Y_2, \dots, Y_n) 是独立同分布随机变量序列, 记其分布为 F_θ . 定义偏经验似然函数

$$\tilde{L}_n(F_\theta) = \prod_{k=1}^n (F_\theta(Y_k) - F_\theta(Y_{k-})) = \prod_{k=1}^n p_k.$$

在只有 $p_k \geq 0$ 和 $\sum_{k=1}^n p_k = 1$ 的约束下, 偏经验似然函数 $\tilde{L}_n(F_\theta)$ 在 $p_k = 1/n (k = 1, \dots, n)$ 处达到最大. 若 OU 型过程从平稳分布出发, 记 X_0 的密度函数为 $p_{\tilde{\theta}} = p_{\tilde{\theta}}(x)$, 则有经验似然比函数

$$R_n(\theta) = p_{\tilde{\theta}} \prod_{k=1}^n n p_k.$$

最大经验似然估计 $\hat{\theta}$ 就是在约束

$$\mathcal{R}_n = \left\{ p_{\tilde{\theta}} = p_{\tilde{\theta}}(X_0); p_k \geq 0 (k = 1, \dots, n), \sum_{k=1}^n p_k = 1; \right. \\ \left. \sum_{k=1}^n p_k g(X_k, X_{k-1}, \theta, s_{l_2}, z_{l_1}) = 0, l_1 = 1, \dots, m_1; l_2 = 1, \dots, m_2 \right\}$$

下, 使得偏经验似然比函数 $R_n(\theta)$ 最大的 θ . 平稳分布的密度函数 $p_{\tilde{\theta}}(\cdot)$ 通常没有简单的闭形式, 但是从后文的证明可知, 当样本量足够大时这一项是可以忽略的.

假设凸集 $\mathcal{R}_n(\theta) = \{\sum_{k=1}^n p_k \mathbf{g}(X_k, X_{k-1}, \theta) \mid p_k \geq 0, \sum_{k=1}^n p_k = 1\}$ 包含 0. 求解最大经验似然估计的最大化问题可表示为 Lagrange 形式

$$L_n(\theta) = \log(p_{\tilde{\theta}}(X_0)) + \sum_{k=1}^n \log(n p_k) - \kappa \left[\sum_{k=1}^n p_k - 1 \right] - n \boldsymbol{\lambda}^T \left[\sum_{k=1}^n p_k \mathbf{g}(X_k, X_{k-1}, \theta) \right], \quad (3.1)$$

其中 κ 和 λ 是 Lagrange 乘子. 文献 [19] 详细的介绍了式 (3.1) 的解. 易证对数经验似然函数为

$$l_n(\theta) = \log \left(p_{\hat{\theta}}(X_0) \prod_{k=1}^n n\hat{p}_k \right) = \log(p_{\hat{\theta}}(X_0)) - \sum_{k=1}^n \log[1 + \lambda(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)]. \quad (3.2)$$

由最大化对数似然函数可得未知参数的最大经验似然估计, 即

$$\hat{\theta} = \arg \max_{\theta} \min_{\lambda} \left\{ \log(p_{\hat{\theta}}(X_0)) - \sum_{k=1}^n \log[1 + \lambda^T \mathbf{g}(X_k, X_{k-1}, \theta)] \right\} \quad (3.3)$$

(详见文献 [19]). 式 (3.2) 也可以通过对偶似然的方法得到 (参见文献 [26]).

3.2 相合性和有效性

假设 A1 \mathbf{X} 是严平稳 α 混合的 OU 型过程, 初始条件 X_0 与背景驱动 Lévy 过程 Z 相互独立. 存在 $r \geq 2, \tau > 0$, 使得混合系数 $\{\alpha(u), u \geq 0\}$ 满足 $\sum_{i=0}^{\infty} [\alpha(i\Delta)]^{\tau/(\tau+1)} < \infty$.

假设 A2 $\int_0^{\Delta} e^{-c(\Delta-s)} dZ(cs)$ 的特征函数 $\phi_{\theta}(z)$ 在 Θ 的紧子集 Θ_1 上关于 θ 连续, Θ_1 包含参数真值 θ_0 .

假设 A3 工具函数 $\gamma_{\theta}(s, x) = \overline{\gamma_{\theta}(-s, x)}$ 关于 (θ, s) 连续且有界, $\lim_{c \rightarrow \infty} \gamma_{\theta}(s, X_0)$ 是 X_0 的非退化的函数, 数学期望 $\mathbb{E}_{\theta_0}[\gamma_{\theta_0}(s, X_0)e^{i\tau e^{-c_0\Delta}X_0}] \neq 0$.

由命题 2.4, 若存在 $q > 0$, 使得 \mathbf{X} 的平稳分布 q 阶绝对矩有限, 则 \mathbf{X} 是指数 β 混合的, 从而假设 A1 成立. 由引理 2.1, 若 $Z(1)$ 的累积函数关于 $\tilde{\theta}$ 连续, 则假设 A2 成立. 通常工具函数 $\gamma_{\theta}(s, x) = \exp(isx)$ 满足假设 A3 的要求.

当离散化参数 τ 的数目随着样本量增加而增大时, 如下的定理 3.1–3.3 分别给出了最大经验似然估计的相合性、渐近正态性和渐近有效性.

定理 3.1 设参数真值 $\theta_0 \in \text{Int}(\Theta_1)$. 取 $m_1 \uparrow \infty, m_2 \uparrow \infty$, 且 $m_1 m_2 = m(n) = [n^{\frac{1}{2}-\eta}]$, 其中 $0 < \eta < \frac{1}{2}$. 令 $z_l = K_1 l / m_1$ ($l = 1, \dots, m_1$), $s_l = K_2 l / m_2$ ($l = 1, \dots, m_2$), K_1 和 K_2 是两个很大的常数. 如果假设 A1–A3 成立, 则由式 (3.3) 定义的最大经验似然估计是弱相合估计, 即当 $n \rightarrow \infty$ 时, $\hat{\theta} \xrightarrow{p} \theta_0$.

假设 B1 $\int_0^{\Delta} e^{-c(\Delta-s)} dZ(cs)$ 的特征函数 $\phi_{\theta}(z)$ 在 Θ 的紧子集 Θ_2 上关于 θ 至少二次连续有界可微, Θ_2 包含参数真值 θ_0 .

假设 B2 工具函数 $\gamma_{\theta}(s, x) = \overline{\gamma_{\theta}(-s, x)}$ 关于 (θ, s) 连续且有界, 关于 θ 可微. 数学期望 $\mathbb{E}_{\theta_0}[\gamma_{\theta_0}(s, X_0)e^{ize^{-c_0\Delta}X_0}] \neq 0$, $\lim_{c \rightarrow \infty} \gamma_{\theta}(s, X_0)$ 是 X_0 的非退化函数, 且满足矩条件: 存在 $\tau > 0$, $\mathbb{E}_{\theta_0}|\gamma_{\theta}(s, X_0)X_0^i|^{2+\tau} < \infty$, $\mathbb{E}_{\theta_0}|\frac{\partial}{\partial\theta}\gamma_{\theta}(s, X_0)X_0^{i-1}|^{2+\tau} < \infty$, $\mathbb{E}_{\theta_0}|\frac{\partial^2}{\partial\theta\partial\theta}\gamma_{\theta}(s, X_0)|^{2+\tau} < \infty$, $i = 1, 2$.

由命题 2.4, 如果存在 $\tau > 0$, 使得 $\mathbb{E}_{\theta_0}|X_0|^{4+2\tau} < \infty$, 则严平稳 OU 型过程 \mathbf{X} 是指数 β 混合的, 进一步取 $\gamma_{\theta}(s, x) = \exp(isx)$, 则假设 A1 和 B2 成立.

设未知参数个数为 p . 为方便起见, 定义

$$\begin{aligned} \Phi_{\theta}(X_0) &= (\phi_{\theta}^R(z_1|X_0), \dots, \phi_{\theta}^R(z_{m_1}|X_0), \phi_{\theta}^I(z_1|X_0), \dots, \phi_{\theta}^I(z_{m_1}|X_0))^T, \\ \Gamma_{\theta}(X_0) &= (\gamma_{\theta}^R(s_1, X_0), \dots, \gamma_{\theta}^R(s_{m_2}, X_0), \gamma_{\theta}^I(s_1, X_0), \dots, \gamma_{\theta}^I(s_{m_2}, X_0))^T, \\ B_m(\theta) &= \mathbb{E}_{\theta_0} \left(\Gamma_{\theta} \otimes \frac{\partial \Phi_{\theta}}{\partial \theta} \right), \\ A_m(\theta) &= \mathbb{E}_{\theta} [\mathbf{g}(X_1, X_0, \theta) \mathbf{g}(X_1, X_0, \theta)^T], \end{aligned}$$

其中 Φ_θ 是 $2m_1$ 元向量, Γ_θ 是 $2m_2$ 元向量, $B_m(\theta)$ 是 $4m_1m_2 \times p$ 矩阵, $A_m(\theta)$ 是 $4m_1m_2 \times 4m_1m_2$ 矩阵. 在适当条件下, 有

$$\mathbb{E}_{\theta_0} \left[\frac{\partial \mathbf{g}(X_1, X_0, \theta_0)}{\partial \theta} \right] = B_m(\theta_0).$$

定义 $p \times p$ 矩阵

$$J_K(\theta) = \lim_{n \rightarrow \infty} [B_m(\theta)^T A_m(\theta)^{-1} B_m(\theta)].$$

定理 3.2 设参数真值 $\theta_0 \in \text{Int}(\Theta_2)$. 取 $m_1 \uparrow \infty$, $m_2 \uparrow \infty$, 且 $m_1m_2 = m(n) = [n^{\frac{1}{2}-\eta}]$, 其中 $0 < \eta < \frac{1}{2}$. 令 $z_l = K_1l/m_1$ ($l = 1, \dots, m_1$), $s_l = K_2l/m_2$ ($l = 1, \dots, m_2$), K_1 和 K_2 是很大的常数. 如果假设 A1, B1 和 B2 成立, 则式 (3.3) 定义的最大经验似然估计是渐近正态的, 即当 $n \rightarrow \infty$ 时, $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N_p(0, J_K^{-1}(\theta_0))$.

如果强度参数 c 已知, 仍有类似定理 3.1 和 3.2 的结论.

推论 3.1 记背景驱动 Lévy 过程 Z 的参数真值为 $\tilde{\theta}_0$. 如果强度参数 c 已知, 在定理 3.1 的条件下, $\tilde{\theta}$ 的最大经验似然估计与 $\tilde{\theta}_0$ 弱相合.

推论 3.2 记背景驱动 Lévy 过程 Z 的参数真值为 $\tilde{\theta}_0$. 如果强度参数 c 已知, 在定理 3.1 的条件下, 如果假设 B1 成立, 则 $\tilde{\theta}$ 最大经验似然估计 $\hat{\tilde{\theta}}$ 是渐近正态的, 即 $\sqrt{n}(\hat{\tilde{\theta}} - \tilde{\theta}_0) \xrightarrow{d} N(0, J_K(\tilde{\theta}_0)^{-1})$.

当 $K_1 \rightarrow \infty$, $K_2 \rightarrow \infty$ 时, 研究 $J_K(\theta_0)$ 的渐近性质, 有助于找到最优工具函数. 记 X_0 的状态空间为 E .

假设 C1 转移密度函数 $p_\theta(y; x, t)$ 满足

1. 对任意 $(x, y) \in E^2$, $p_\theta(y; x, t)$ 在 Θ 的紧子集 Θ_3 上关于 θ 二次连续可微, Θ_3 包含参数真值 θ_0 . 对任意 $\theta \in \Theta_3$, 有 $\mathbb{E}_\theta \left[\frac{\partial \log p_\theta(X_1; X_0, \Delta)}{\partial \theta} \middle| X_0 \right] = 0$ a.s.

2. 对任意 $\theta \in \Theta_3$, Fisher 信息阵

$$I(\theta) = \mathbb{E}_\theta \left[\left(\frac{\partial \log p_\theta(X_1; X_0, \Delta)}{\partial \theta} \right) \left(\frac{\partial \log p_\theta(X_1; X_0, \Delta)}{\partial \theta} \right)^T \right]$$

非奇异.

定理 3.3 取 $\gamma_\theta(s, x) = e^{isx}$. 在定理 3.2 的条件下, 如果假设 C1 成立, 令 $K_1 \rightarrow \infty$, $K_2 \rightarrow \infty$, 则由式 (3.3) 定义的最大经验似然估计 $\hat{\theta}$ 是渐近有效的, 即

$$\lim_{\substack{K_1 \rightarrow \infty \\ K_2 \rightarrow \infty}} J_K(\theta_0) = I(\theta_0).$$

采用单指标工具函数的最大经验似然估计有类似的渐近性质. 取 $m = m(n) = [n^{\frac{1}{2}-\eta}]$ ($0 < \eta < 1/2$), 在 m 个不同点上计算条件特征函数和工具函数, 离散点为 $z = z_l$, $l = 1, 2, \dots, m$, 且 $z_1 < z_2 < \dots < z_m$. 记 $4m$ 元向量

$$\mathbf{g}^{SI}(X_k, X_{k-1}, \theta) = (g(X_k, X_{k-1}, \theta, z_1, z_1)^T, \dots, g(X_k, X_{k-1}, \theta, z_m, z_m)^T)^T.$$

定义 $B_m^{SI}(\theta) = \mathbb{E}_{\theta_0} \left[\frac{\partial \mathbf{g}^{SI}(X_1, X_0, \theta)}{\partial \theta} \right]$, $A_m^{SI}(\theta) = \mathbb{E}_{\theta_0} [\mathbf{g}^{SI}(X_1, X_0, \theta) \mathbf{g}^{SI}(X_1, X_0, \theta)^T]$, $J_K^{SI}(\theta) = \lim_{n \rightarrow \infty} [B_m^{SI}(\theta)^T A_m^{SI}(\theta)^{-1} B_m^{SI}(\theta)]$. 如下的定理 3.4 和 3.5 给出了基于单指标工具变量的最大经验似然估计的渐近性质.

定理 3.4 取 $m = m(n) = [n^{\frac{1}{2}-\eta}]$, $z_l = Kl/m$ ($l = 1, \dots, m$), 其中 $0 < \eta < \frac{1}{2}$, K 是很大的常数. 在定理 3.1 的条件下, 基于单指标工具变量的最大经验似然估计是参数 θ 的弱相

合估计. 在定理 3.2 的条件下, 基于单指标工具变量的最大经验似然估计渐近于均值为 0, 方差为 $J_K^{SI}(\theta_0)^{-1}$ 的正态分布.

定理 3.5 取 $m = m(n) = [n^{\frac{1}{2}-\eta}]$, $z_l = Kl/m$ ($l = 1, \dots, m$), 其中 $0 < \eta < \frac{1}{2}$, K 是很大的常数. 如果工具函数 $\gamma_\theta(z, x) = \frac{1}{2\pi} \int \frac{\partial \log p_\theta(y; x, \Delta)}{\partial \theta} e^{-izx} dy$ 满足假设 B2, 在定理 3.3 的条件下, 基于单指标工具变量的最大经验似然估计是渐近有效的, 即 $\lim_{K \rightarrow \infty} J_K^{SI}(\theta_0) = I(\theta_0)$.

如果强度参数 c 已知, 则引理 3.1 中的序列 $\{Y_k = X_k - e^{-c\Delta} X_{k-1}, k \in \mathbb{N}\}$ 可以观测到, 前述定理的结论与 Kunitomo 和 Owada^[20] 的类似, 容易验证 $\gamma_\theta(z, x) = e^{-ize^{-c\Delta}x}$ 也是最优单指标工具函数.

假设 C1 要求转移密度函数关于强度参数 c 可微, 但在很多情形下不满足. 当背景驱动 Lévy 过程 Z 是 A 型或 B 型时 (参见文献 [21, 定义 11.19]), $\int_0^t e^{-c(t-s)} dZ(cs)$ 的密度函数通常可表示为

$$f_\theta(x, t) = (d_\theta)^t \delta(x) \mathbb{I}_{\{x=0\}} + \bar{f}_\theta(x, t) \mathbb{I}_{\{x \neq 0\}}. \quad (3.4)$$

其中 $\delta(x)$ 表示在 x 点具有概率质量 (参见文献 [21]). 例如, 对于参数为 (c, a, b) 的 Gamma-OU 过程, 转移密度函数 $p_\theta(y; x, t) = f_\theta(y - e^{-ct}x, t)$ 关于 c 存在不连续点. Fisher 信息阵 $I(\theta_0)$ 没有定义, 可以找到更有效的估计. 这种情形下, 如果样本量足够大, 就有可能准确的找到强度参数 c 的真值.

引理 3.2 如果 $\int_0^t e^{-c(t-s)} dZ(cs)$ 的密度函数是式 (3.4) 的形式, 且 $d_\theta > 0$, $\bar{f}_\theta(\cdot, t)$ 连续. 记 X_M 是样本序列 $\{X_{k-1}/X_k, k = 1, \dots, n\}$ 的众数, 则条件概率 $P(\ln(X_M)/\Delta \neq c \mid \#\{X_M\} \geq 2) = 0$, 概率 $P(\#\{X_M\} \geq 2) = 1 - (1 - d_\theta)^n - nd_\theta(1 - d_\theta)^{n-1}$, 其中 $\#\{X_M\}$ 是 X_M 在序列中出现的次数.

基于此, 可按以下步骤对式 (3.3) 定义的最大似然估计进行修正:

1. 如果 $\#\{X_M\} \geq 2$, 则令 $\hat{c}_M = \ln(X_M)/\Delta$. 以 \hat{c}_M 为强度参数, 采用工具函数 $\gamma(s, x) = e^{isx}$ 或 $\gamma(z, x) = e^{-ize^{-c\Delta}x}$ 得 $\tilde{\theta}$ 的最大经验似然估计 $\hat{\theta}_M$. 最后取 $\hat{\theta}_M = (\hat{c}_M, \hat{\theta}_M)$.
2. 否则不修正, $\hat{\theta}_M$ 直接取为式 (3.3) 定义的最大经验估计.

随机模拟的结果有力的支持了引理 3.3 的结论. 通常当样本量较大时, 在第 1 步中就能找出强度参数 c 的真值, 而不需要转入第 2 步.

定理 3.6 如果 $\int_0^t e^{-c(t-s)} dZ(cs)$ 的密度函数有式 (3.4) 的形式, 并且 $d_\theta > 0$, $\bar{f}_\theta(\cdot, t)$ 连续.

1. 在定理 3.1 的条件下, $\hat{\theta}_M$ 是 θ 的弱相合估计.
2. 在定理 3.2 的条件下, $\hat{\theta}_M$ 渐近正态, 特别的, $\hat{\theta}_M$ 是渐近有效的, 即 $\lim_{n \rightarrow \infty} P(\hat{c} \neq c_0) = 0$, 渐近方差 $A\text{Var}(\hat{c}) = 0$, $A\text{Var}(\hat{\theta}) = I(\tilde{\theta}_0)$.

4 经验似然比检验

如下的定理 4.1 和 4.2 给出参数检验问题和估计方程检验问题的经验似然比检验方法.

定理 4.1 在定理 3.2 的条件下, 参数检验问题 $H_0 : \theta = \theta_0$ 的经验似然比统计量为 $W_1 = 2[l_n(\hat{\theta}_n) - l_n(\theta_0)]$, 其中对数经验似然函数 $l_n(\theta)$ 由式 (3.2) 定义, 则当 H_0 为真, 样本量 $n \rightarrow \infty$ 时, 有 $W_1 \xrightarrow{d} \chi^2(p)$.

定理 4.2 在定理 3.2 的条件下, 估计方程的检验问题 $H_0 : E_{\theta_0}[g(X_1, X_0, \theta_0)] = 0$ 的经

验似然比统计量为 $W_2 = -2l_n(\hat{\theta}_n)$, 其中对数经验似然函数 $l_n(\theta)$ 由式 (3.2) 定义, 则当 H_0 为真, 样本量 $n \rightarrow \infty$ 时,

$$\frac{W_2 - 4m_1m_2}{\sqrt{16m_1m_2}} \xrightarrow{d} N(0, 1).$$

5 随机模拟

Owen 在文献 [19] 中详细的介绍了最大化式 (3.2) 定义的对数经验似然函数 $l_n(\theta)$ 的数值方法. 二步最优化方法是常用方法. 本文模拟中最大化方法由统计软件 R 中的 “optim” 函数及其 “L-BFGS-B” 的最优化方法实现, 并采用样本简单矩估计作为参数 $\tilde{\theta}$ 的初值, 对数一阶经验自回归函数作为强度参数 c 的初值. 模拟实现了第 2 节中介绍的两类 OU 型过程: Gamma-OU 过程和 IG-OU 过程. 参数估计的每种情形都做了 1000 次模拟, 检验问题的每种情形都做了 3000 次模拟. 观测时间间隔 $\Delta = 1$ 固定. 在 $m_1 = 5$ 个点上对条件特征函数进行离散化, 取常数 $K_1 = 45$. 为了更好的利用 0 附近的信息, 离散点取为 $t = lK_1/m_1 - 8.9$ ($l = 1, \dots, 5$), 即 $t = (0.1, 9.1, 18.1, 27.1, 36.1)$. 取 $m_2 = K_2 = 1$, 工具函数的离散点 $s = m_2/K_2 = 1$.

对 Gamma-OU 过程的模拟中, 参数空间取为 $a \in [0.001, 1000]$, $b \in [0.001, 1000]$, $c \in [0.001, 1000]$. 平稳分布的均值和方差分别为 a/b 和 a/b^2 , 初值选为 $\hat{a}_0 = \bar{X}^2/S_X^2$, $\hat{b}_0 = \bar{X}/S_X^2$, 其中 $\bar{X} = \frac{1}{n} \sum_{k=0}^n X_k$, $S_X^2 = \frac{1}{n-1} \sum_{k=0}^n (X_k - \bar{X})^2$. 模拟结果见表 1, 估计值的下方括号中是均方根误差 (RMSE). 因所有模拟中得到的都是 c 的真值 (引理 3.3), 因此没有列在表中.

表 1 Gamma-OU 过程参数的最大经验似然估计

True		n=200		n=500		n=1000		
c	a	b	\hat{a}	\hat{b}	\hat{a}	\hat{b}	\hat{a}	\hat{b}
0.6	3	3	3.0782 (0.4314)	3.0886 (0.4146)	3.0020 (0.2432)	3.0063 (0.2350)	3.0253 (0.1577)	3.0183 (0.1474)
0.6	3	1.2	3.0548 (0.4113)	1.2273 (0.1617)	3.0451 (0.2546)	1.2213 (0.0939)	3.0305 (0.1878)	1.2162 (0.0730)

本文构造的最大经验似然估计的均方根误差 (RMSE) 比表 2 中的最大似然估计 (参见文献 [6]) 的均方根误差稍大, 但是当样本量增大时, 差距快速缩小, 最大经验似然估计得到很大改善. 而且当 $b \in (0, 2^{\frac{2}{3}})$ 时, 最大似然估计的性质未知 (参见文献 [6]), 因此无法保证最大似然估计的效果. 但最大经验似然估计没有这个问题.

表 2 Gamma-OU 过程参数的最大似然估计

True		n=200		n=500		
c	a	b	\hat{a}	\hat{b}	\hat{a}	\hat{b}
0.6	3	3	3.0047 (0.2538)	3.0114 (0.2607)	2.9976 (0.1484)	3.0092 (0.1514)

参数估计和估计方程问题的 CP 值 (coverage probability) 见表 3. $CP(W_1)$ 是参数 (a, b) 的 95% 置信区间的 CP 值, 模拟结果说明经验似然比统计量 W_1 具有较好的性质, 当样本量

在 300 和 500 之间的时候, 就有接近 95% 的 CP 值. 但是对估计方程的检验问题而言, 经验似然比统计量 W_2 (渐近正态) 还需进一步改进.

表 3 Gamma-OU 过程经验似然比统计量的 CP 值

True			$n=200$		$n=300$		$n=500$		$n=1000$	
c	a	b	CP(W_1)	CP(W_2)	CP(W_1)	CP(W_2)	CP(W_1)	CP(W_2)	CP(W_1)	CP(W_2)
0.6	3	3	93.4%	97.9%	94.9%	97.8%	95.3%	97.0%	95.3%	96.7%
		6	91.4%	97.6%	94.5%	97.4%	95.4%	97.4%	95.5%	97.7%

参数 (a, b) 的 95% 置信区间的模拟见图 1.

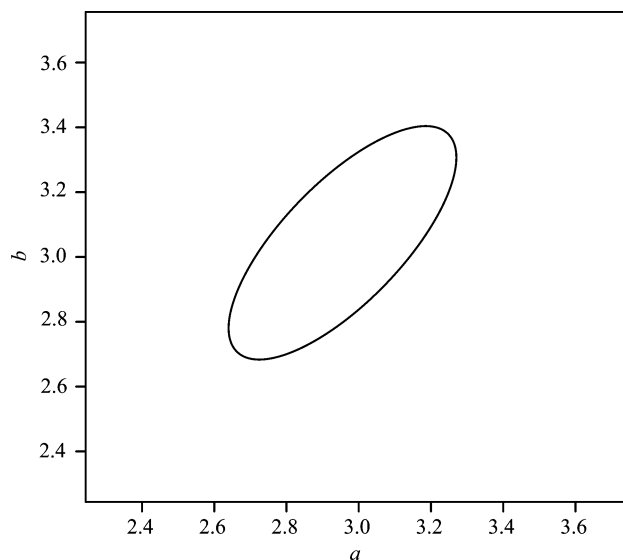


图 1 参数 (a, b) 的 95% 置信区间

本文对检验问题的势作了模拟研究. 参数检验的势的模拟结果见表 4. 取 $c = 0.6$, 原假设 $H_0 : a = 3, b = 3$, 备择假设 $H_1 : a = a_1, b = b_1$, 例如 $H_1 : a = 3, b = 2.8$. 参数检验的势随着样本量的增加而增大, 也随着备择假设与原假设的距离增加而增大.

表 4 Gamma-OU 过程 W_1 的势

(a_1, b_1)	$n = 200$	$n = 300$	$n = 500$	$n = 1000$
(3, 2.8)	0.124	0.152	0.220	0.417
(3.2, 3.2)	0.131	0.144	0.164	0.257
(3, 2.7)	0.229	0.322	0.513	0.803
(3.3, 3.3)	0.210	0.220	0.281	0.495
(3, 2.5)	0.608	0.775	0.938	0.999
(3.5, 3.5)	0.370	0.455	0.648	0.901

在估计方程检验的势的模拟研究中, 取原假设为: 样本来自 Gamma-OU 过程, 估计方程基于 Gamma-OU 过程的条件特征函数构造的. Gamma-OU 过程的背景驱动 Lévy 过程是一个复合 Poisson 过程, 该过程的跳服从指数分布. 备择假设选择了两种情形进行比较. 第 1 种

情形, 备择假设为真时, 背景驱动 Lévy 过程的跳服从对数正态分布 ($\mu = 1, \sigma^2 = 1$). 第 2 种情形, 备择假设为真时, 背景驱动 Lévy 过程的跳服从卡方分布 ($df = 1$). 模拟结果见表 5.

表 5 Gamma-OU 过程 W_2 的势

	$n = 200$	$n = 300$	$n = 500$	$n = 1000$
log-N(1,1)	0.035	0.049	0.115	0.330
$\chi^2(1)$	0.035	0.078	0.222	0.735

IG-OU 过程是 Tempered-stable OU 过程和 Normal-inverse-Gaussian OU 过程的一种特殊情形, 经常用来刻画波动率. 它不符合引理 3.3 的条件, 无法找出强度参数 c 的真值. 我们取参数空间 $a \in [0.001, 1000]$, $b \in [0.001, 1000]$, $c \in [0.001, 1000]$. 平稳分布的均值和方差分别为 a/b 和 a/b^2 , 初值选为 $\hat{a}_0 = \bar{X}^2/S_X^2$ 和 $\hat{b}_0 = \bar{X}/S_X^2$. 模拟结果见表 6, 估计值下方括号中的是均方根误差 (RMSE).

表 6 IG-OU 过程参数的最大经验似然估计

True			$n=200$			$n=500$		
c	a	b	\hat{a}	\hat{b}	\hat{c}	\hat{a}	\hat{b}	\hat{c}
0.5	1	1	1.0464	1.0754	0.5237	1.0145	1.0293	0.5027
			(0.1174)	(0.1761)	(0.1027)	(0.0557)	(1.0980)	(0.0076)
	10	2	10.1479	(2.0327)	0.5332	10.0525	2.0131	0.5138
			(0.7977)	(0.1662)	(0.1337)	(0.5324)	(0.1186)	(0.0965)

上述模拟结果显示基于条件特征函数的最大经验似然估计相当准确稳定, 特别在样本量较大的情形下有较大优势. 直接利用条件特征函数构造矩条件避免 Fourier 逆变换等数值运算, 具有计算的便利性. 而且在许多情形下, 条件特征函数有更明确的闭形式, 易于实现对数经验似然函数的最优化.

6 结论

本文将经验似然方法应用于离散观测 OU 型过程的参数估计和检验问题, 研究了最大经验似然估计和经验似然比统计量的性质. 基于条件特征函数的最大经验似然估计具有相合性和渐近正态性, 如果采用最优工具函数, 所得最大经验似然估计还是渐近有效的, 所以最大经验似然估计具有重要的实用价值, 特别对处理金融市场中的高频数据有很大优势. 本文给出的经验似然方法也可以推广到一般离散观测随机过程的参数估计问题, 其主要思想是构造参数或半参数的估计方程, 找出最优工具函数, 得到渐近有效的经验似然估计.

7 证明

引理 2.1 的证明 该结论可由文献 [21] 的引理 17.1 和式 (2.1) 直接得到.

固定 $t > 0$, 先求随机变量 $Y(0, t)$ 的分布是重要的. 由文献 [21] 中引理 17.1, 随机变量 $Y(0, t)$ 具有如下特征函数:

$$\mathbb{E}[e^{izY(0,t)}] = \exp \left\{ c \int_0^t \psi(e^{-cs} z) ds \right\},$$

其中 $\psi(z)$ 是 $Z(1)$ 的累积函数.

引理 3.1 的证明 该结论可由引理 2.1 及 Lévy 过程的独立增量性直接得到.

引理 7.1 如果存在正的常数 Γ , 使得 $|\gamma_\theta(s, x)| < \Gamma$ 对任意 $s \in \mathbb{R}$ 和 $x \in E$ 都成立, 则由式 (3.1) 定义的约束方程 \mathbf{g} 有界, 即对任意 $(x, y) \in E^2$ 和 $\theta \in \Theta$ 都有

$$\|\mathbf{g}(y, x, \theta)\| \leq 4\Gamma\sqrt{m_1m_2}.$$

引理 7.1 的证明是直接的.

定理 3.1 的证明 首先引入两个常用记号: $O_p(\cdot)$ 和 $o_p(\cdot)$. 随机变量序列 Y_n 和 Z_n , 如果 Y_n/Z_n 依概率有界, 则记 $Y_n = O_p(Z_n)$; 如果 Y_n/Z_n 依概率收敛到 0, 则记 $Y_n = o_p(Z_n)$. 容易验证 $\frac{1}{n} \log(p_\theta(X_0)) = O_p(\frac{1}{n})$, 这一项在后边的证明里微不足道, 可以忽略.

$\text{NB}(\theta_0, \delta)$ 是 θ_0 的半径为 δ 的邻域. 定义准则函数

$$G_n(\theta) = -\frac{1}{n} \sum_{k=1}^n \log[1 + K_1K_2\boldsymbol{\lambda}(\theta)^\top \mathbf{g}(X_k, X_{k-1}, \theta)], \quad (7.1)$$

这里为了强调常数 K_1K_2 在渐近有效性中的作用, 将 $\boldsymbol{\lambda}(\theta)$ 乘以 K_1K_2 , 并不影响结论. 整个证明由两部分组成, 即证明对任一 $\delta > 0$, 存在 $d(\delta) < 0$, 使得当 $n \rightarrow \infty$ 时, 分别成立

$$\mathbb{P}\left(\sup_{\theta^* \in \Theta \setminus \text{NB}(\theta_0, \delta)} \sqrt{\frac{n}{m_1m_2}} G_n(\theta^*) > d(\delta)\right) < \frac{\delta}{2}, \quad (7.2)$$

和

$$\mathbb{P}\left(\sqrt{\frac{n}{m_1m_2}} G_n(\hat{\theta}) < d(\delta)\right) < \frac{\delta}{2}. \quad (7.3)$$

首先找到使得式 (7.2) 成立的 $d(\delta)$. 因为 OU 型过程 \mathbf{X} 是严平稳的, $\{\mathbf{g}(X_k, X_{k-1}, \theta), k = 1, \dots, n\}$ 是一独立同分布变量序列. 定义两个辅助函数

$$\mathbf{u}(\theta) = \frac{(m_1m_2)^{-\frac{1}{2}}K_1K_2\mathbb{E}_{\theta_0}[\mathbf{g}(X_1, X_0, \theta)]}{1 + \sqrt{\frac{K_1K_2}{m_1m_2}}\|\mathbb{E}_{\theta_0}[\mathbf{g}(X_1, X_0, \theta)]\|},$$

$$v_k(\theta) = -\log[1 + n^{-\frac{1}{2}}\mathbf{u}(\theta)^\top \mathbf{g}(X_k, X_{k-1}, \theta)] \quad (k = 1, \dots, n).$$

由 Taylor 定理, 存在 $t \in (0, 1)$, 使得

$$\begin{aligned} & \sup_{\theta \in \Theta_1} \left| \sqrt{\frac{n}{m_1m_2}} \frac{1}{n} \sum_{k=1}^n v_k(\theta) - \frac{1}{\sqrt{m_1m_2}n} \sum_{k=1}^n [-\mathbf{u}(\theta)^\top \mathbf{g}(X_k, X_{k-1}, \theta)] \right| \\ & \leq \sup_{\theta \in \Theta_1} \left| \frac{1}{n} \sum_{k=1}^n \frac{(m_1m_2n)^{-\frac{1}{2}}[\mathbf{u}(\theta)^\top \mathbf{g}(X_1, X_0, \theta)]^2}{2[1 + n^{-\frac{1}{2}}t\mathbf{u}(\theta)^\top \mathbf{g}(X_1, X_0, \theta)]^2} \right|. \end{aligned} \quad (7.4)$$

由引理 (7.1) 知

$$\sup_{\theta \in \Theta_1} n^{-\frac{1}{2}}\mathbf{u}(\theta)^\top \mathbf{g}(X_1, X_0, \theta) \leq 16K_1K_2\Gamma^2\sqrt{m_1m_2/n},$$

及

$$\sup_{\theta \in \Theta_1} (m_1m_2n)^{-\frac{1}{2}}[\mathbf{u}(\theta)^\top \mathbf{g}(X_1, X_0, \theta)]^2 \leq 16(K_1K_2)^2\Gamma^4\sqrt{m_1m_2/n}.$$

当 $n \rightarrow \infty$ 时, 两者都趋向于 0, 所以不等式 (7.4) 右边趋向于 0.

注意到对任意 $(x_0, x_1) \in E^2$ 和 $\theta \in \Theta_1$ 都有,

$$\begin{aligned} \mu(x_1, x_0, \theta) & \stackrel{\text{def}}{=} \lim_{m_1 \rightarrow \infty} \frac{K_1K_2}{m_1m_2} [\mathbb{E}_{\theta_0} \mathbf{g}(X_1, X_0, \theta)]^\top \mathbf{g}(x_1, x_0, \theta) \\ & = \int_0^{K_1} \int_0^{K_2} [\mathbb{E}_{\theta_0} \mathbf{g}(X_1, X_0, \theta, s, z)]^\top \mathbf{g}(x_1, x_0, \theta, s, z) ds dz \leq 16K_1K_2\Gamma^2, \end{aligned}$$

$$\begin{aligned}\mu_\theta &\stackrel{\text{def}}{=} \lim_{\substack{m_1 \rightarrow \infty \\ m_2 \rightarrow \infty}} \frac{K_1 K_2}{m_1 m_2} \|\mathbb{E}_{\theta_0} \mathbf{g}(X_1, X_0, \theta)\|^2 \\ &= \int_0^{K_1} \int_0^{K_2} \|\mathbb{E}_{\theta_0} \mathbf{g}(X_1, X_0, \theta, s, z)\|^2 ds dz \leq 16 K_1 K_2 \Gamma^2.\end{aligned}$$

在假设 A2 和 A3 的条件下, $g(x_1, x_0, \theta, s, z)$ 有界, 且关于 (θ, s, z) 在紧集 $\Theta_1 \times [0, K_2] \times [0, K_1]$ 上一致连续. 据 Arzelà-Ascoli 定理,

$$\lim_{\substack{m_1 \rightarrow \infty \\ m_2 \rightarrow \infty}} \sup_{\theta \in \Theta_1} \left| \frac{K_1 K_2}{m_1 m_2} [\mathbb{E}_{\theta_0} \mathbf{g}(X_1, X_0, \theta)]^T \mathbf{g}(x_1, x_0, \theta) - \mu(x_1, x_0, \theta) \right| = 0.$$

类似的,

$$\lim_{\substack{m_1 \rightarrow \infty \\ m_2 \rightarrow \infty}} \sup_{\theta \in \Theta_1} \left| \sqrt{\frac{K_1 K_2}{m_1 m_2}} \|\mathbb{E}_{\theta_0} \mathbf{g}(X_1, X_0, \theta)\| - \sqrt{\mu_\theta} \right| = 0.$$

因此,

$$\begin{aligned}\sup_{\theta \in \Theta_1} &\left| \frac{1}{\sqrt{m_1 m_2}} \mathbf{u}(\theta)^T \mathbf{g}(x_1, x_0, \theta) - \frac{\mu(x_1, x_0, \theta)}{1 + \sqrt{\mu_\theta}} \right| \\ &\leq \sup_{\theta \in \Theta_1} \left| \frac{K_1 K_2}{m_1 m_2} [\mathbb{E}_{\theta_0} \mathbf{g}(X_1, X_0, \theta)]^T \mathbf{g}(x_1, x_0, \theta) - \mu(x_1, x_0, \theta) \right| \\ &\quad + \sup_{\theta \in \Theta_1} |\mu(x_1, x_0, \theta)| \cdot \sup_{\theta \in \Theta_1} \left| \sqrt{\frac{K_1 K_2}{m_1 m_2}} \|\mathbb{E}_{\theta_0} \mathbf{g}(X_1, X_0, \theta)\| - \sqrt{\mu_\theta} \right| \rightarrow 0.\end{aligned}\quad (7.5)$$

式 (7.5) 的左边有界, 因此

$$\mathbb{E}_{\theta_0} \sup_{\theta \in \Theta_1} \left| \frac{1}{\sqrt{m_1 m_2}} \mathbf{u}(\theta)^T \mathbf{g}(X_1, X_0, \theta) - \frac{\mu(X_1, X_0, \theta)}{1 + \sqrt{\mu_\theta}} \right|^2 \rightarrow 0.$$

由 Markov 不等式, 对任意 $\epsilon > 0$,

$$\begin{aligned}\mathbb{P} \left(\sup_{\theta \in \Theta_1} \left| \frac{1}{n} \sum_{k=1}^n \left[\frac{1}{\sqrt{m_1 m_2}} \mathbf{u}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta) - \frac{\mu(X_k, X_{k-1}, \theta)}{1 + \sqrt{\mu_\theta}} \right] \right| > \frac{\epsilon}{2} \right) \\ \leq \frac{8}{\epsilon^2} \mathbb{E}_{\theta_0} \sup_{\theta \in \Theta_1} \left| \frac{1}{\sqrt{m_1 m_2}} \mathbf{u}(\theta)^T \mathbf{g}(X_1, X_0, \theta) - \frac{\mu(X_1, X_0, \theta)}{1 + \sqrt{\mu_\theta}} \right|^2 \rightarrow 0.\end{aligned}\quad (7.6)$$

令 $U(x_1, x_0, \theta) = -\mu(x_1, x_0, \theta)/(1 + \sqrt{\mu_\theta})$. 在假设 A2 和 A3 的条件下, $U(x_1, x_0, \theta)$ 有界. 又在假设 A1 的条件下, \mathbf{X} 是 α 混合的. 由 Billingsley 不等式,

$$\text{Cov}(U(X_k, X_{k-1}, \theta), U(X_l, X_{l-1}, \theta)) \leq 4\alpha_{kl} \|U(X_1, X_0, \theta)\|_\infty^2.$$

所以存在常数 d_1 只与混合系数 $\{\alpha(t)\}$ 有关, 使得

$$\text{Var}_{\theta_0} \left(\frac{1}{n} \sum_{k=1}^n U(X_k, X_{k-1}, \theta) \right) \leq \frac{d_1}{n} \|U(X_1, X_0, \theta)\|_\infty^2 = o(1).$$

由 Markov 不等式,

$$\mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n U(X_k, X_{k-1}, \theta) > \frac{-\mu_\theta}{1 + \sqrt{\mu_\theta}} + \frac{\epsilon}{2} \right) \leq \frac{\text{Var}_{\theta_0} \left(\frac{1}{n} \sum_{k=1}^n U(X_k, X_{k-1}, \theta) \right)}{\left[\frac{-\mu_\theta}{1 + \sqrt{\mu_\theta}} + \frac{\epsilon}{2} - \mathbb{E}_{\theta_0} U(X_1, X_0, \theta) \right]^2} = o(1).$$

由 α 混合过程的一致大数定律, 知当 $n \rightarrow \infty$ 时,

$$\mathbb{P} \left(\sup_{\theta^* \in \Theta_1 \setminus \text{NB}(\theta_0, \delta)} \frac{1}{n} \sum_{k=1}^n U(X_k, X_{k-1}, \theta^*, \rho_{\theta^*}) > \sup_{\theta^* \in \Theta_1 \setminus \text{NB}(\theta_0, \delta)} \frac{-\mu_{\theta^*}}{1 + \sqrt{\mu_{\theta^*}}} + \frac{\epsilon}{2} \right) \rightarrow 0, \quad (7.7)$$

上式对所有 $\epsilon > 0$ 都成立. 令

$$d(\delta) = \frac{1}{2} \sup_{\theta^* \in \Theta_1 \setminus \text{NB}(\theta_0, \delta)} \frac{-\mu_{\theta^*}}{1 + \sqrt{\mu_{\theta^*}}}.$$

显然 $d(\delta) < 0$. 综合式 (7.4), (7.6) 和 (7.7) 知, 存在 $\bar{n}(\delta)$, 对任意 $n \geq \bar{n}(\delta)$, 有

$$\begin{aligned} & \mathbb{P} \left(\sup_{\theta^* \in \Theta_1 \setminus \text{NB}(\theta_0, \delta)} \sqrt{\frac{n}{m_1 m_2}} \frac{1}{n} \sum_{k=1}^n v_k(\theta^*) > d(\delta) \right) \\ & \leq \mathbb{P} \left(\sup_{\theta^* \in \Theta_1 \setminus \text{NB}(\theta_0, \delta)} \frac{1}{\sqrt{m_1 m_2} n} \sum_{k=1}^n [-\mathbf{u}(\theta^*)^\top \mathbf{g}(X_k, X_{k-1}, \theta^*)] > 2d(\delta) \right) \\ & \leq \mathbb{P} \left(\sup_{\theta^* \in \Theta_1 \setminus \text{NB}(\theta_0, \delta)} \left| \frac{1}{n} \sum_{k=1}^n \left[\frac{1}{\sqrt{m_1 m_2}} \mathbf{u}(\theta^*)^\top \mathbf{g}(X_k, X_{k-1}, \theta^*) - \frac{\mu(X_k, X_{k-1}, \theta^*)}{1 + \sqrt{\mu_{\theta^*}}} \right] \right| > -d(\delta) \right) \\ & \quad + \mathbb{P} \left(\sup_{\theta^* \in \Theta_1 \setminus \text{NB}(\theta_0, \delta)} \frac{1}{n} \sum_{k=1}^n \frac{-\mu(X_k, X_{k-1}, \theta^*)}{1 + \sqrt{\mu_{\theta^*}}} > d(\delta) \right) \\ & < \frac{\delta}{2}. \end{aligned}$$

从而成立

$$\mathbb{P} \left(\sup_{\theta^* \in \Theta_1 \setminus \text{NB}(\theta_0, \delta)} \sqrt{\frac{n}{m_1 m_2}} G_n(\theta^*) > d(\delta) \right) < \frac{\delta}{2},$$

现在只需证明式 (7.3) 成立. Lagrange 乘子是方程组

$$\frac{1}{n} \sum_{k=1}^n \frac{K_1 K_2 \mathbf{g}(X_k, X_{k-1}, \theta)}{1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^\top \mathbf{g}(X_k, X_{k-1}, \theta)} = 0 \quad (7.8)$$

的解. 记 $\boldsymbol{\lambda}(\theta_0) = \|\boldsymbol{\lambda}(\theta_0)\| \boldsymbol{\xi}$, 其中 $\boldsymbol{\xi}$ 是单位向量, 则

$$\begin{aligned} 0 &= \frac{1}{m_1 m_2} \left\| \frac{1}{n} \sum_{k=1}^n \frac{K_1 K_2 \mathbf{g}(X_k, X_{k-1}, \theta_0)}{1 + K_1 K_2 \boldsymbol{\lambda}(\theta_0)^\top \mathbf{g}(X_k, X_{k-1}, \theta_0)} \right\| \\ &\geq \frac{\|\boldsymbol{\lambda}(\theta_0)\|}{1 + K_1 K_2 \|\boldsymbol{\lambda}(\theta_0)^\top\| \max_{1 \leq i \leq n} \|\mathbf{g}(X_k, X_{k-1}, \theta_0)\|} \\ &\quad \times \left[\frac{1}{n} \sum_{k=1}^n \frac{(K_1 K_2)^2}{m_1 m_2} \boldsymbol{\xi}^\top \mathbf{g}(X_k, X_{k-1}, \theta_0) \mathbf{g}(X_k, X_{k-1}, \theta_0)^\top \boldsymbol{\xi} \right] \\ &\quad - \left| \frac{1}{n} \sum_{k=1}^n \frac{K_1 K_2}{m_1 m_2} \boldsymbol{\xi}^\top \mathbf{g}(X_k, X_{k-1}, \theta_0) \right|. \end{aligned} \quad (7.9)$$

在假设 A1 和 A3 的条件下, 由 Hölder's 不等式和 Billingsley 不等式, 存在只与混合系数 $\{\alpha(t)\}$ 有关的常数 $d_2 > 0$, 使得

$$\mathbb{E}_{\theta_0} \left[\left| \frac{1}{n} \sum_{k=1}^n \boldsymbol{\xi}^\top \mathbf{g}(X_k, X_{k-1}, \theta_0) \right|^2 \right] \leq \frac{16d_2 \Gamma m_1 m_2}{n} = O\left(\frac{m_1 m_2}{n}\right).$$

因此 $\left| \frac{1}{n} \sum_{k=1}^n \boldsymbol{\xi}^\top \mathbf{g}(X_k, X_{k-1}, \theta_0) \right| = O_p(\sqrt{m_1 m_2/n})$. 定义 $4m_1 m_2 \times 4m_1 m_2$ 矩阵 $D(n) = (D_{ij}(n))$ 如下:

$$\begin{aligned} D(n) &= \frac{1}{n} \sum_{k=1}^n \frac{(K_1 K_2)^2}{m_1 m_2} \{ \mathbf{g}(X_k, X_{k-1}, \theta_0) \mathbf{g}(X_k, X_{k-1}, \theta_0)^\top \\ &\quad - \mathbb{E}_{\theta_0} [\mathbf{g}(X_k, X_{k-1}, \theta_0) \mathbf{g}(X_k, X_{k-1}, \theta_0)^\top] \}. \end{aligned} \quad (7.10)$$

由 Markov 不等式和 Billingsley 不等式, 对任意 $\epsilon > 0$, 都有 $P(|D_{ij}(n)| \geq \epsilon) \leq d_3/(nm_1^2m_2^2)$, 及 $P(\max_{1 \leq i, j \leq 4m_1m_2} |D_{ij}(n)| \geq \epsilon) \leq d_4/n$, 其中 c_3 和 c_4 是正的常数. 记

$$\Sigma_m(\theta_0) = \frac{(K_1K_2)^2}{m_1m_2} \mathbb{E}_{\theta_0}[\mathbf{g}(X_k, X_{k-1}, \theta_0)\mathbf{g}(X_k, X_{k-1}, \theta_0)^\top].$$

根据文献 [20] 的讨论, $\Sigma_m(\theta_0)$ 正定, 特征根有界. 由式 (7.9), 知

$$\frac{\|\boldsymbol{\lambda}(\theta_0)\|}{1 + K_1K_2\|\boldsymbol{\lambda}(\theta_0)\| \max_{1 \leq k \leq n} \|\mathbf{g}(X_k, X_{k-1}, \theta_0)\|} = O_p\left(\sqrt{\frac{1}{m_1m_2n}}\right).$$

据引理 7.1, $\max_{1 \leq k \leq n} \|\mathbf{g}(X_k, X_{k-1}, \theta_0)\| \leq 4\Gamma\sqrt{m_1m_2} = O(\sqrt{m_1m_2})$. 因此,

$$\|\boldsymbol{\lambda}(\theta_0)\| = O_p\left(\sqrt{\frac{1}{m_1m_2n}}\right). \quad (7.11)$$

由于 $\log(1+x) \leq x$,

$$G_n(\theta_0) \geq -\frac{1}{n} \sum_{k=1}^n K_1K_2\boldsymbol{\lambda}(\theta_0)^\top \mathbf{g}(X_k, X_{k-1}, \theta_0) = O_p\left(\frac{1}{n}\right).$$

所以存在 $n(\delta) \in \mathcal{N}$, 对任意 $n \geq n(\delta)$, 都有

$$P\left(\sqrt{\frac{n}{m_1m_2}}G_n(\theta_0) \leq d(\delta)\right) < \frac{\delta}{2}.$$

由最大经验似然估计的定义 (3.3), 知 $G_n(\hat{\theta}) = \sup_{\theta^* \in \Theta_1} G_n(\theta^*)$, 及

$$P\left(\sqrt{\frac{n}{m_1m_2}}G_n(\hat{\theta}) \leq d(\delta)\right) < \frac{\delta}{2}.$$

综合式 (7.2) 和 (7.3), 可知

$$P(\hat{\theta} \notin \text{NB}(\theta_0, \delta)) < \delta.$$

定理得证.

定理 3.2 的证明 一阶条件 $\frac{\partial G_n(\hat{\theta})}{\partial \theta} = 0$. 由中值定理, $\frac{\partial G_n(\hat{\theta})}{\partial \theta}$ 在 $\hat{\theta} = \theta_0$ 处展开, 得

$$-\sqrt{n}\frac{\partial G_n(\theta_0)}{\partial \theta} = \frac{\partial^2 G_n(\hat{\theta}^*)}{\partial \theta \partial \theta} \sqrt{n}(\hat{\theta}_n - \theta_0), \quad (7.12)$$

其中 $\|\hat{\theta}^* - \hat{\theta}_n\| \leq \|\theta^* - \theta_0\|$. 定理的证明分两部分, 分别证明

$$-\sqrt{n}\frac{\partial G_n(\theta_0)}{\partial \theta} \xrightarrow{d} N_p(0, J_K(\theta_0)), \quad (7.13)$$

及

$$\frac{\partial^2 G_n(\theta^*)}{\partial \theta \partial \theta} = -J_K(\theta_0) + o_p(1), \quad (7.14)$$

其中 $J_K(\theta_0)$ 非奇异.

式 (7.12) 的左边可展开为

$$-\sqrt{n}\frac{\partial G_n(\theta_0)}{\partial \theta} = \sqrt{n}\frac{1}{n} \sum_{k=1}^n \frac{K_1K_2}{1 + K_1K_2\boldsymbol{\lambda}(\theta_0)^\top \mathbf{g}(X_k, X_{k-1}, \theta_0)} \left(\frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta_0)}{\partial \theta}\right)^\top \boldsymbol{\lambda}(\theta_0), \quad (7.15)$$

其中

$$\frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta)}{\partial \theta} = \left(\frac{\partial \Gamma_\theta(X_{k-1})}{\partial \theta} \otimes \mathbf{h}(X_k, X_{k-1}, \theta) - \Gamma_\theta(X_{k-1}) \otimes \frac{\partial \Phi_\theta(X_{k-1})}{\partial \theta}\right),$$

$\Gamma_\theta(X_{k-1})$, $\Phi_\theta(X_{k-1})$ 和 $\mathbf{h}(X_k, X_{k-1}, \theta)$ 如前定义.

$\lambda(\theta_0)$ 是方程组 (7.8) 的解. 由方程组 (7.8),

$$0 = \frac{1}{n} \sum_{k=1}^n \frac{K_1 K_2}{m_1 m_2} \mathbf{g}(X_k, X_{k-1}, \theta_0) - \left\{ \frac{1}{n} \sum_{k=1}^n \frac{(K_1 K_2)^2}{m_1 m_2} \mathbf{g}(X_k, X_{k-1}, \theta_0) \mathbf{g}(X_k, X_{k-1}, \theta_0)^\top \right\} \lambda(\theta_0) + r_{1n},$$

易证余项 $\|r_{1n}\| = o_p(1)$. 式 (7.12) 的左边可以近似为

$$-\sqrt{n} \frac{\partial G_n(\theta_0)}{\partial \theta} = \sqrt{n} \left\{ \frac{1}{n} \sum_{k=1}^n \frac{K_1 K_2}{1 + K_1 K_2 \lambda(\theta_0)^\top \mathbf{g}(X_k, X_{k-1}, \theta_0)} \left(\frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta_0)}{\partial \theta} \right)^\top \right\} \times S_{n, \theta_0}^{-1} \left\{ \frac{1}{n} \sum_{k=1}^n \frac{K_1 K_2}{m_1 m_2} \mathbf{g}(X_k, X_{k-1}, \theta_0) \right\} + r_{2n},$$

其中

$$S_{n, \theta} = \frac{1}{n} \sum_{k=1}^n \frac{(K_1 K_2)^2}{m_1 m_2} \mathbf{g}(X_k, X_{k-1}, \theta_0) \mathbf{g}(X_k, X_{k-1}, \theta_0)^\top,$$

易证余项 $\|r_{2n}\| = o_p(1)$. 式 (7.12) 的左边可进一步近似为

$$-\sqrt{n} \frac{\partial G_n(\theta_0)}{\partial \theta} = \frac{1}{\sqrt{n}} \sum_{k=1}^n K_1 K_2 B_m(\theta_0)^\top \{ \mathbb{E}_{\theta_0} S_{n, \theta_0} \}^{-1} \frac{K_1 K_2}{m_1 m_2} \mathbf{g}(X_k, X_{k-1}, \theta_0) + o_p(1).$$

对任意 $\zeta \in \mathbb{R}^p$, 定义随机变量阵列 $\{Y_{mk}, k = 1, \dots, n\}$ 如下:

$$Y_{mi} = \zeta' K_1 K_2 B_m(\theta_0)^\top \{ \mathbb{E}_{\theta_0} (S_{n, \theta_0}) \}^{-1} \frac{K_1 K_2}{m_1 m_2} \mathbf{g}(X_k, X_{k-1}, \theta_0).$$

在假设 A1 的条件下, $\{Y_{mk}, k = 1, \dots, n\}$ 严平稳, α 混合. 其方差为

$$\text{Var}_{\theta_0}(Y_{m1}) = \frac{(K_1 K_2)^2}{m_1 m_2} \zeta' B_m(\theta_0)^\top \{ \mathbb{E}_{\theta_0} (S_{n, \theta_0}) \}^{-1} B_m(\theta_0) \zeta.$$

由 Billingsley 不等式, 知

$$\frac{n \text{Var}_{\theta_0}(Y_{m1})}{\text{Var}_{\theta_0}(\sum_{k=1}^n Y_{mk})} = O(1).$$

对任意 $\kappa > 0$,

$$\begin{aligned} & \frac{\sum_{k=1}^n \mathbb{E}_{\theta_0} [Y_{mk}^2 : |Y_{mk}| > \kappa (\text{Var}_{\theta_0}(\sum_{k=1}^n Y_{mk}))^{1/2}]}{\sum_{k=1}^n \text{Var}_{\theta_0}(Y_{mk})} \\ &= \frac{\mathbb{E}_{\theta_0} \left[\left(\frac{Y_{m1}}{\sqrt{m_1 m_2}} \right)^2 m_1 m_2 \mathbb{I} \left(\left| \frac{Y_{m1}}{\sqrt{m_1 m_2}} \right| > \kappa \left(\frac{n}{m_1 m_2} \text{Var}_{\theta_0}(Y_{m1}) \right)^{1/2} \right) \right]}{\text{Var}_{\theta_0}(Y_{m1})} \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

据 α 混合序列的 Lindeberg 型中心极限定理 (参见文献 [25]), 对任意 $\zeta \in \mathbb{R}^p$, 都有

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n \zeta' K_1 K_2 B_m(\theta_0)^\top \{ \mathbb{E}_{\theta_0} (S_{n, \theta_0}) \}^{-1} \frac{K_1 K_2}{m_1 m_2} \mathbf{g}(X_k, X_{k-1}, \theta_0) \xrightarrow{d} N(\mathbf{0}, \zeta' J_K(\theta_0) \zeta).$$

因此

$$-\sqrt{n} \frac{\partial G_n(\theta_0)}{\partial \theta} \xrightarrow{d} N_p(0, J_K(\theta_0)).$$

二阶导数可展开为

$$\begin{aligned} \frac{\partial^2 G_n(\theta)}{\partial \theta \partial \theta} &= \frac{(K_1 K_2)^2}{n} \sum_{k=1}^n \frac{\left(\frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta)}{\partial \theta} \right)^T \boldsymbol{\lambda}(\theta) \boldsymbol{\lambda}(\theta)^T \left(\frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta)}{\partial \theta} \right)}{[1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)]^2} \\ &+ \frac{(K_1 K_2)^2}{n} \sum_{k=1}^n \frac{\left(\frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta)}{\partial \theta} \right)^T \boldsymbol{\lambda}(\theta) \mathbf{g}(X_k, X_{k-1}, \theta)^T \left(\frac{\partial \boldsymbol{\lambda}(\theta)}{\partial \theta} \right)}{[1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)]^2} \\ &- \frac{K_1 K_2}{n} \sum_{k=1}^n \frac{\left(\frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta)}{\partial \theta} \right)^T \left(\frac{\partial \boldsymbol{\lambda}(\theta)}{\partial \theta} \right)}{1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)} \\ &- \frac{K_1 K_2}{n} \sum_{k=1}^n \sum_{l=1}^4 \sum_{j=1}^{m_1 m_2} \lambda_{lj}(\theta) \frac{\left(\frac{\partial^2 g_{(l-1)m_1 m_2 + j}(X_k, X_{k-1}, \theta)}{\partial \theta \partial \theta} \right)}{1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)}, \end{aligned} \quad (7.16)$$

其中 $\mathbf{g}(X_k, X_{k-1}, \theta) = (g_i(X_k, X_{k-1}, \theta))$ ($k = 1, \dots, n$).

记 $\text{NB}(\theta_0) = \{\theta \mid \|\theta - \theta_0\| \leq n^{\frac{1}{3}}\}$. Lagrange 乘子 $\boldsymbol{\lambda}(\theta)$ 满足,

$$\frac{1}{n} \sum_{k=1}^n \frac{K_1 K_2 \mathbf{g}(X_k, X_{k-1}, \theta)}{1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)} = 0. \quad (7.17)$$

在假设 B1 和 B2 的条件下, 由 Markov 不等式和 Davydov 不等式, 知式 (7.16) 等号右边的第 1 项和第 4 项都是 $o_p(1)$. 第 2 项和第 3 项中包含 θ 的导数. 对式 (7.17) 关于 θ 求导得

$$T_n(\theta) \frac{\partial \boldsymbol{\lambda}(\theta)}{\partial \theta} = \frac{1}{n} \sum_{k=1}^n \frac{K_1 K_2}{m_1 m_2} \frac{\frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta)}{\partial \theta}}{1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)} - R_n(\theta),$$

其中

$$T_n(\theta) = \frac{1}{n} \sum_{k=1}^n \frac{(K_1 K_2)^2}{m_1 m_2} \frac{\mathbf{g}(X_k, X_{k-1}, \theta) \mathbf{g}(X_k, X_{k-1}, \theta)^T}{[1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)]^2},$$

及

$$R_n(\theta) = \frac{1}{n} \sum_{k=1}^n \frac{(K_1 K_2)^2}{m_1 m_2} \frac{\mathbf{g}(X_k, X_{k-1}, \theta) \boldsymbol{\lambda}(\theta)^T \frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta)}{\partial \theta}}{[1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)]^2}.$$

与式 (7.10) 的讨论类似, 有

$$T_n(\theta) - \frac{(K_1 K_2)^2}{m_1 m_2} A_m(\theta) = o_p(1),$$

及 $R_n(\theta) = o_p(1)$. 式 (7.16) 等号右边的第 3 项可近似为

$$\begin{aligned} &\left\{ \frac{K_1 K_2}{n} \sum_{k=1}^n \frac{\left(\frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta)}{\partial \theta} \right)^T}{1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)} \right\} T_n(\theta)^{-1} \\ &\times \left\{ \frac{K_1 K_2}{n} \sum_{k=1}^n \frac{K_1 K_2}{m_1 m_2} \frac{\frac{\partial \mathbf{g}(X_k, X_{k-1}, \theta)}{\partial \theta}}{1 + K_1 K_2 \boldsymbol{\lambda}(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)} \right\} + o_p(1). \end{aligned}$$

定理得证.

定理 3.3 的证明 为方便起见, 定义复向量

$$\begin{aligned} \tilde{\Phi}_\theta &= (\phi_\theta(z_1|X_0), \dots, \phi_\theta(z_{m_1}|X_0), \phi_\theta(-z_1|X_0), \dots, \phi_\theta(-z_{m_1}|X_0))^T, \\ \tilde{\Gamma}_\theta &= (e^{is_1 X_0}, \dots, e^{is_{m_2} X_0}, e^{-is_1 X_0}, \dots, e^{-is_{m_2} X_0})^T, \end{aligned}$$

$$\begin{aligned}\tilde{\mathbf{h}}(X_1, X_0, \theta) &= (e^{iz_1 X_1} - \phi_\theta(z_1|X_0), \dots, e^{iz_{m_1} X_1} - \phi_\theta(z_{m_1}|X_0), \\ &\quad e^{-iz_1 X_1} - \phi_\theta(-z_1|X_0), \dots, e^{-iz_{m_1} X_1} - \phi_\theta(-z_{m_1}|X_0))^T, \\ \tilde{\mathbf{g}}(X_1, X_0, \theta) &= \tilde{\Gamma}_\theta \otimes \tilde{\mathbf{h}}(X_1, X_0, \theta).\end{aligned}$$

易证

$$\begin{aligned}B_m(\theta_0)^T A_m(\theta_0)^{-1} B_m(\theta_0) &= \left[\mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \right]^T \\ &\quad \times [\mathbb{E}_{\theta_0} \tilde{\mathbf{g}}(X_1, X_0, \theta_0) \tilde{\mathbf{g}}(X_1, X_0, \theta_0)^T]^{-1} \mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right).\end{aligned}$$

令

$$w_\theta(s, z) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\partial \log p_\theta(y; x, \Delta)}{\partial \theta} e^{-i(sx+zy)} dx dy,$$

其中 $p_\theta(y; x, \Delta)$ 是转移密度函数. 定义 $4m_1 m_2 \times p$ 矩阵

$$\tilde{W}_\theta = (w_\theta(s_1, z_1), \dots, w_\theta(s_1, z_{m_1}), \dots, w_\theta(s_{m_2}, z_1), \dots, w_\theta(s_{m_2}, z_{m_1}))^T.$$

则当 $n \rightarrow \infty$ ($m_1, m_2 \rightarrow \infty$) 时, 有极限

$$\frac{K_1 K_2}{m_1 m_2} \tilde{W}_{\theta_0}^T \mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \rightarrow \int_{-K_2}^{K_2} \int_{-K_1}^{K_1} w_{\theta_0}(s, z) \mathbb{E}_{\theta_0} \left(e^{isX_0} \frac{\partial \tilde{\phi}_{\theta_0}(z|X_0)}{\partial \theta} \right)^T dz ds.$$

据 Fourier 逆变换公式, 得

$$\frac{1}{2\pi} \int e^{-isx} \left(\int e^{isu} \frac{\partial \phi_\theta(z|u)}{\partial \theta} p_{\tilde{\theta}}(u) du \right) ds = \frac{\partial \phi_\theta(z|x)}{\partial \theta} p_{\tilde{\theta}}(x),$$

其中 $p_{\tilde{\theta}}(x)$ 是边际密度函数, 又

$$\frac{1}{2\pi} \int \frac{\partial \phi_\theta(z|x)}{\partial \theta} e^{-izy} dz = \frac{\partial p_\theta(y; x, \Delta)}{\partial \theta}.$$

联合密度函数可表示为 $p_{\theta_0}(x_1; x_0, \Delta) p_{\tilde{\theta}_0}(x_0)$, 若求导和积分运算可交换, 则

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w_{\theta_0}(s, z) \mathbb{E}_{\theta_0} \left(e^{isX_0} \frac{\partial \tilde{\phi}_{\theta_0}(z|X_0)}{\partial \theta} \right)^T dz ds = I(\theta_0).$$

用类似于 Feuerverger 和 McDunnough^[9] 的方法, 可得

$$\lim_{\substack{K_1 \rightarrow \infty \\ K_2 \rightarrow \infty}} \lim_{n \rightarrow \infty} \frac{K_1 K_2}{m_1 m_2} \tilde{W}_{\theta_0}^T \left[\mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \right] = I(\theta_0).$$

类似的, 有

$$\lim_{\substack{K_1 \rightarrow \infty \\ K_2 \rightarrow \infty}} \lim_{n \rightarrow \infty} \left(\frac{K_1 K_2}{m_1 m_2} \right)^2 \tilde{W}_{\theta_0}^T \mathbb{E}_{\theta_0} [\tilde{\mathbf{g}}(X_1, X_0, \theta_0) \tilde{\mathbf{g}}(X_1, X_0, \theta_0)^T] \tilde{W}_{\theta_0} = I(\theta_0).$$

对任意非零 $4m_1 m_2 \times p$ 矩阵 V ,

$$\begin{aligned}&\left\{ \left[V^T \mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \right]^T [V^T \mathbb{E}_{\theta_0} \tilde{\mathbf{g}}(X_1, X_0, \theta_0) \tilde{\mathbf{g}}(X_1, X_0, \theta_0)^T V]^{-1} \right. \\ &\quad \left. \times \left[V^T \mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \right]^T \right\}^{-1}\end{aligned}$$

在

$$V = [\mathbb{E}_{\theta_0} \tilde{\mathbf{g}}(X_1, X_0, \theta_0) \tilde{\mathbf{g}}(X_1, X_0, \theta_0)^T]^{-1} \mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right)$$

处最小化, 最小值为

$$\left\{ \left[\mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \right]^T \left[\mathbb{E}_{\theta_0} \tilde{\mathbf{g}}(X_1, X_0, \theta_0) \tilde{\mathbf{g}}(X_1, X_0, \theta_0)^T \right]^{-1} \mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \right\}^{-1}.$$

因 Fisher 信息阵 $I(\theta_0)$ 非奇异, 所以有

$$\begin{aligned} V^T I(\theta_0)^{-1} V &\leq V^T \left\{ \left[\mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \right]^T \left[\mathbb{E}_{\theta_0} \tilde{\mathbf{g}}(X_1, X_0, \theta_0) \tilde{\mathbf{g}}(X_1, X_0, \theta_0)^T \right]^{-1} \right. \\ &\quad \left. \times \mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \right\}^{-1} V \\ &\leq V^T \left\{ \left[\tilde{W}_{\theta_0}^T \mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \right]^T \right. \\ &\quad \left. \times (\tilde{W}_{\theta_0}^T \mathbb{E}_{\theta_0} [\tilde{\mathbf{g}}(X_1, X_0, \theta_0) \tilde{\mathbf{g}}(X_1, X_0, \theta_0)^T] \tilde{W}_{\theta_0})^{-1} \right. \\ &\quad \left. \times \left[\tilde{W}_{\theta_0}^T \mathbb{E}_{\theta_0} \left(\tilde{\Gamma}_{\theta_0} \otimes \frac{\partial \tilde{\Phi}_{\theta_0}}{\partial \theta} \right) \right] \right\}^{-1} V. \end{aligned}$$

定理得证.

定理 3.4 的证明 单指标最大经验似然估计的相合性的证明与定理 3.1 类似, 渐近正态性的证明与定理 3.2 类似, 此处从略.

定理 3.5 的证明与定理 3.3 类似, 此处从略.

引理 3.2 的证明 若存在 $1 \leq i < j \leq n$, 使得 $\log(X_{i-1}/X_i) = \log(X_{j-1}/X_j) \neq c\Delta$. 因 $\bar{f}_\theta(\cdot, t)$ 连续, 所以

$$P(\log(X_{i-1}/X_i) = \log(X_i/X_{i+1}) \neq c\Delta) = 0,$$

及

$$\begin{aligned} &P(\ln(X_M)/\Delta \neq c, \#\{X_M\} \geq 2) \\ &= P(\text{存在 } 1 \leq i < j \leq n \text{ 使得 } \log(X_{i-1}/X_i) = \log(X_{j-1}/X_j) \neq c\Delta) = 0. \end{aligned}$$

可知

$$\begin{aligned} &P(\#\{X_M\} \geq 2) \\ &= P(\text{存在 } 1 \leq i < j \leq n \text{ 使得 } x_i - e^{-c\Delta} x_{i-1} = x_i - e^{-c\Delta} x_{i-1} = 0) \\ &= 1 - (1 - d_\theta)^n - nd_\theta(1 - d_\theta)^{n-1}. \end{aligned}$$

因此 $P(\ln(X_M)/\Delta \neq c \mid \#\{X_M\} \geq 2) = 0$. 引理得证.

定理 3.6 的证明 该结论可由引理 3.3 和定理 3.3 直接得到.

定理 4.1 的证明 定理的证明方法类似于 Kunitomo 和 Owada^[20] 的证明, 但在本文模型中, OU 型过程是弱相依的, 观测样本间的相关性不容忽视. 由于证明所需的矩条件在定理 3.1 及 3.2 的证明中都已经给出, 本证明中不再赘述. 本证明主要针对采用双指标工具变量的经验似然比统计量, 采用单指标工具变量的经验似然比统计量性质的证明与之类似.

令 $Y_k(\theta) = K_1 K_2 \lambda(\theta)^T \mathbf{g}(X_k, X_{k-1}, \theta)$ ($k = 1, \dots, n$). 准则函数 $G_n(\theta)$ 可在 $\theta = \hat{\theta}_n$ 展开为

$$G_n(\hat{\theta}_n) = -\frac{1}{n} \sum_{k=1}^n \left[Y_k(\hat{\theta}_n) - \frac{1}{2} Y_k(\hat{\theta}_n)^2 + \frac{1}{3} Y_k(\theta^*)^3 \right],$$

其中 $\|\theta^* - \hat{\theta}_n\| \leq \|\hat{\theta}_n - \theta_0\|$,

$$\begin{aligned} \sum_{k=1}^n Y_k(\hat{\theta}_n) &= K_1 K_2 \sum_{k=1}^n \boldsymbol{\lambda}(\theta_0)^\top \mathbf{g}(X_k, X_{k-1}, \theta_0) \\ &\quad - \sqrt{n}(\hat{\theta}_n - \theta_0)^\top B_m(\theta_0)^\top A_m(\theta_0)^{-1} B_m(\theta_0)(\hat{\theta}_n - \theta_0)\sqrt{n} + o_p(1), \\ \sum_{k=1}^n Y_k(\hat{\theta}_n)^2 &= (K_1 K_2)^2 \sum_{k=1}^n [\boldsymbol{\lambda}(\theta_0)^\top \mathbf{g}(X_k, X_{k-1}, \theta_0)]^2 \\ &\quad - \sqrt{n}(\hat{\theta}_n - \theta_0)^\top B_m(\theta_0)^\top A_m(\theta_0)^{-1} B_m(\theta_0)(\hat{\theta}_n - \theta_0)\sqrt{n} + o_p(1). \\ \sum_{k=1}^n |Y_k(\theta^*)|^3 &\leq (K_1 K_2)^2 \left[\max_{1 \leq k \leq n} |Y_k(\theta^*)| \right] \sum_{k=1}^n |\boldsymbol{\lambda}(\theta^*)^\top \mathbf{g}(X_k, X_{k-1}, \theta^*)|^2 \\ &\leq K_1 K_2 (m_1 m_2)^{\frac{3}{2}} \|\boldsymbol{\lambda}(\theta^*)\| [\sqrt{n} \boldsymbol{\lambda}(\theta^*)]^\top \\ &\quad \times \left[\frac{(K_1 K_2)^2}{n} \sum_{k=1}^n \frac{1}{m_1 m_2} \mathbf{g}(X_k, X_{k-1}, \theta^*) \mathbf{g}(X_k, X_{k-1}, \theta^*)^\top \right] [\sqrt{n} \boldsymbol{\lambda}(\theta^*)] \\ &= o_p(1), \end{aligned}$$

类似地, $G_n(\theta_0)$ 可近似为

$$\begin{aligned} -2nG_n(\theta_0) &= K_1 K_2 \sum_{k=1}^n \boldsymbol{\lambda}(\theta_0)^\top \mathbf{g}(X_k, X_{k-1}, \theta_0) \\ &\quad - \frac{1}{2} (K_1 K_2)^2 \sum_{k=1}^n [\boldsymbol{\lambda}(\theta_0)^\top \mathbf{g}(X_k, X_{k-1}, \theta_0)]^2 + o_p(1), \end{aligned}$$

从而有

$$2nG_n(\hat{\theta}_n) = 2nG_n(\theta_0) + \sqrt{n}(\hat{\theta}_n - \theta_0)^\top B_m(\theta_0)^\top A_m(\theta_0)^{-1} B_m(\theta_0)(\hat{\theta}_n - \theta_0)\sqrt{n} + o_p(1).$$

据定理 3.2 中最大经验似然估计的渐近正态性, 定理得证.

定理 4.2 的证明 与定理 4.1 的证明类似, 我们不再赘述前述证明中的一些中间结论.

令

$$\begin{aligned} \Sigma_m(\theta_0) &= (\sigma_{jk}(m)) = \frac{(K_1 K_2)^2}{m_1 m_2} \mathbb{E}_{\theta_0} [\mathbf{g}(X_k, X_{k-1}, \theta_0) \mathbf{g}(X_k, X_{k-1}, \theta_0)^\top], \\ \Sigma_m(\theta_0)^{-1} &= (\sigma^{jk}(m)). \end{aligned}$$

显然

$$\sum_{j,k=1}^{4m_1 m_2} \sigma^{jk}(m) \sigma_{jk}(m) = 4m_1 m_2.$$

记

$$\mathbf{Y}_n = (Y_{nj}) = \frac{1}{\sqrt{n}} \sum_{k=1}^n \frac{K}{\sqrt{m}} \mathbf{g}(X_k, X_{k-1}, \theta_0).$$

类似于定理 4.1 证明中的方法, 检验统计量 W_2 可近似为 $\mathbf{Y}_n^\top \Sigma_m(\theta_0)^{-1} \mathbf{Y}_n + O_p(1)$. 对主要项有

$$\frac{\mathbf{Y}_n^\top \Sigma_m(\theta_0)^{-1} \mathbf{Y}_n - 4m_1 m_2}{\sqrt{16m_1 m_2}}$$

$$\begin{aligned}
&= \frac{1}{\sqrt{16m_1m_2}} \frac{1}{n} \sum_{i=i'=1}^n \sum_{j,k=1}^{4m_1m_2} \left[\frac{(K_1K_2)^2}{m_1m_2} g_j(X_i, X_{i-1}, \theta_0) g_k(X_{i'}, X_{i'-1}, \theta_0) \sigma^{jk}(m) \right. \\
&\quad \left. - \sigma^{jk}(m) \sigma_{jk}(m) \right] \\
&\quad + \frac{1}{\sqrt{16m_1m_2}} \frac{1}{n} \sum_{i=i'=1}^n \sum_{j,k=1}^{4m_1m_2} \left[\frac{(K_1K_2)^2}{m_1m_2} g_j(X_i, X_{i-1}, \theta_0) g_k(X_{i'}, X_{i'-1}, \theta_0) \sigma^{jk} \right].
\end{aligned}$$

等号右边的数学期望为 0, 且等号右边每项都有界. 由 Billingsley 不等式, 第 1 项的方差小于

$$\frac{1}{n^2} \frac{1}{16m_1m_2} (4m_1m_2)^2 = \frac{m_1m_2}{n^2},$$

当 $n \rightarrow \infty$ 时, 趋向于 0. 对第 2 项可应用 α 混合序列的中心极限定理, 则定理得证.

致谢 作者诚挚的感谢两位审稿人提出的宝贵修改意见. 作者对张世斌博士和王柯博士在论文写作过程中提供的帮助表示感谢.

参考文献

- 1 Barndorff-Nielsen O E, Shephard N. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion). *J Roy Statist Soc Ser B*, **63**: 167–241 (2001)
- 2 Lo A W. Maximum likelihood estimation of generalized Itô processes with discretely sampled data. *Econometric Theory*, **4**: 231–247 (1988)
- 3 Pedersen A R. A new approach to maximum-likelihood estimation for stochastic differential equations based on discrete observations. *Scand J Statist*, **22**: 55–61 (1995)
- 4 Santa-Clara P. Simulated likelihood estimation of diffusions with an application to the short term interest rate. PhD Dissertation. Los Angeles: University of California, 1995
- 5 Barndorff-Nielsen O E. Processes of normal inverse Gaussian type. *Finance Stoch*, **2**: 41–68 (1997)
- 6 Zhang S, Zhang X, Sun S. Parametric estimation of discretely sampled Gamma-OU processes. *Sci China Ser A-Math*, **49**: 1231–1257 (2006)
- 7 Valdivieso L, Schoutens W, Tuerlinckx F. Maximum likelihood estimation in processes of Ornstein-Uhlenbeck type. DOI: 10.1007/S11203-008-9021-8. *Stat Infer Stoch Process*, to appear
- 8 Jongbloed G, van der Meulen F H, van der Vaart A W. Nonparametric inference for Lévy-driven Ornstein-Uhlenbeck processes. *Bernoulli*, **11**: 759–791 (2005)
- 9 Feuerverger A, McDunnough P. On the efficiency of empirical characteristic function procedures. *J Roy Statist Soc Ser B*, **43**: 20–27 (1981)
- 10 Feuerverger A. An efficiency result for the empirical characteristic function in stationary time-series models. *Canad J Statist*, **18**: 155–161 (1990)
- 11 Singleton K J. Estimation of affine asset pricing models using the empirical characteristic function. *J Econometrics*, **102**: 111–141 (2001)
- 12 Jiang G J, Knight J L. Estimation of continuous-time processes via the empirical characteristic function. *J Bus Econom Statist*, **20**: 198–212 (2002)
- 13 Chacko G, Viceira L M. Spectral GMM estimation of continuous-time processes. *J Econometrics*, **116**: 259–292 (2003)
- 14 Owen A B. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, **75**: 237–249 (1988)
- 15 Owen A B. Empirical likelihood ratio confidence regions. *Ann Statist*, **18**: 90–120 (1990)
- 16 Qin J, Lawless J. Empirical likelihood and general estimating equations. *Ann Statist*, **22**: 300–325 (1994)

- 17 Wang Q, Linton O, Hardle W. Semiparametric regression analysis with missing response at random. *J Amer Statist Assoc*, **99**: 334–345 (2004)
- 18 Zhu L, Xue L. Empirical likelihood confidence regions in a partially linear single-index model. *J Roy Statist Soc Ser B*, **68**: 549–570 (2006)
- 19 Owen A B. Empirical Likelihood. London: Chapman and Hall, 2001
- 20 Kunitomo N, Owada T. Empirical likelihood estimation of Lévy processes. Discussion Paper CIRJE-F-272, Graduate School of Economics, University of Tokyo, 2006
- 21 Sato K. Lévy Processes and Infinitely Divisible Distributions. Cambridge: Cambridge University Press, 1999
- 22 Bosq D. Nonparametric Statistics for Stochastic Processes. New York: Springer-Verlag, 1998
- 23 Rosenblatt M. Asymptotic normality, strong mixing and spectral density estimates. *Ann Probab*, **12**: 1167–1180 (1984)
- 24 Masuda H. On multidimensional Ornstein-Uhlenbeck processes driven by a general Lévy process. *Bernoulli*, **10**: 97–120 (2004)
- 25 Doukhan P. Mixing: Properties and Examples. Berlin: Springer-Verlag, 1995
- 26 Mykland P A. Dual likelihood. *Ann Statist*, **23**: 396–421 (1995)