Adaptive maximum-likelihood-type estimation for discretely observed diffusion processes with observational noise

Shogo H. Nakakita[1] and Masayuki Uchida[1,2]

[1]Graduate School of Engineering Science, Osaka University
[2]Center for Mathematical Modeling and Data Science, Osaka University

---

We consider a d-dimensional ergodic diffusion process defined by the following stochastic differential equation such that

$$\mathrm{d}X_t = b\left(X_t,\beta\right)\mathrm{d}t + a\left(X_t,\alpha\right)\mathrm{d}w_t,\ X_0 = x_0,$$

where $\{w_t\}_{t\geq 0}$ is an r-dimensional Wiener process, x_0 is a random variable independent of $\{w_t\}_{t\geq 0}$, $\alpha\in\Theta_1$ and $\beta\in\Theta_2$ are unknown parameters, $\Theta_1\subset\mathbf{R}^{m_1}$ and $\Theta_2\subset\mathbf{R}^{m_2}$ are bounded and open sets in $\mathbf{R}^{m_i}$ for $i=1,2$, $\theta^\star=\left(\alpha^\star,\beta^\star\right)$ is the true value of the parameter, and $a:\mathbf{R}^d\times\Theta_1\to\mathbf{R}^d\otimes\mathbf{R}^r$ and $b:\mathbf{R}^d\times\Theta_2\to\mathbf{R}^d$ are known functions.

Our interest is to examine adaptive maximum-likelihood-type estimators of the diffusion parameter α and the drift one β for discretely observed diffusion processes with observational noise. The discrete and contaminated observation is defined as the sequence of random variables $\left\{Y_{ih_n}\right\}_{i=0,\ldots,n}$ such that for all $i=0,\ldots,n$,

$$Y_{ih_n} = X_{ih_n} + \Lambda^{1/2}\varepsilon_{ih_n},$$

where $h_n>0$ is the discretisation step such that $h_n\to 0$ and $T_n := nh_n\to\infty$ as $n\to\infty$, $\left\{\varepsilon_{ih_n}\right\}_{i=0,\ldots,n}$ is the i.i.d. sequence of d-dimensional random variables independent of $\{X_t\}_{t\geq 0}$ such that $\mathbf{E}\left[\varepsilon_0\right]=\mathbf{0}$ and $\mathrm{Var}\left(\varepsilon_0\right)=I_d$, and Λ is the positive semi-definite matrix determining the variance of the noise term.

To extract the state of the latent process, we compose the sequence of local means with respect to the k_n partitions, that is, $\bar{Y}_j := \frac{1}{p_n}\sum_{i=0}^{p_n-1} Y_{j\Delta_n+ih_n}$, where $p_n := n/k_n$ indicates the number of observation in each partition, and $\Delta_n := p_n h_n$ is the bandwidth of the partitions. With this sequence, we construct two quasi-likelihood functions for these parameters such that

$$\mathbb{H}_{1,n}^{\tau}\left(\alpha;\Lambda\right) := -\frac{1}{2}\sum_{j=1}^{k_n-2}\left(\left(\frac{2}{3}\Delta_n A_n^{\tau}\left(\bar{Y}_{j-1},\alpha,\Lambda\right)\right)^{-1}\left[\left(\bar{Y}_{j+1}-\bar{Y}_j\right)^{\otimes 2}\right] + \log\det A_n^{\tau}\left(\bar{Y}_{j-1},\alpha,\Lambda\right)\right),$$

$$\mathbb{H}_{2,n}\left(\beta;\alpha\right) := -\frac{1}{2}\sum_{j=1}^{k_n-2}\left(\left(\Delta_n A\left(\bar{Y}_{j-1},\alpha\right)\right)^{-1}\left[\left(\bar{Y}_{j+1}-\bar{Y}_j-\Delta_n b\left(\bar{Y}_{j-1},\beta\right)\right)^{\otimes 2}\right]\right),$$

where for all matrices M, M^T is the transpose of M and $M^{\otimes 2} := MM^T$, for all the set of matrices M_1 and M_2 with the same size, $M_1\left[M_2\right] := \mathrm{tr}\left\{M_1 M_2^T\right\}$, $A\left(x,\alpha\right)=a^{\otimes 2}\left(x,\alpha\right)$, $A_n^{\tau}\left(x,\alpha,\Lambda\right) := A\left(x,\alpha\right)+3\Delta_n^{\frac{2-\tau}{\tau-1}}\Lambda$ and $\tau\in(1,2]$ is a tuning parameter.

It is possible to optimise these quasi-likelihoods separately, and this adaptive procedure has an advantage over the existent simultaneous one (e.g., Favetto (2014) and Favetto (2016)) in computational burden. In this talk, we show that both the estimators have asymptotic properties such as consistency, asymptotic normality with different convergence rate and asymptotic independence (for details, see Nakakita and Uchida (2017:arXiv:1711.04462, 2018:arXiv:1805.11414)).