Statistical Inference in partial observation setting, in continuous time

Alexandre Brouste
(with M. Kleptsyna and A. Popier)

March 23rd, 2011

SAPS VIII, Le Mans

PARTIAL OBSERVATION SETTING

In various fields, the *signal* process $X = (X_t, t \geq 0)$, valued in $\mathbb{R}$ and whose law depends on an unknown parameter $\vartheta \in \Theta$ (bounded open subset of $\mathbb{R}$), can not be observed directly but only through an *observation* process

$$Y = (Y_t, t \geq 0) \ .$$

In classical setting, the observation process is given by

$$Y_t = Y_0 + \int_0^t h(X_s)\, ds + \sigma W_t, \quad t \geq 0$$

where $h(.)$ is a known function and $\sigma > 0$ a known constant, $W = (W_t, t \geq 0)$ is a Wiener process (or any martingale) independent of X, Y_0 is a the initial condition (say $Y_0 = 0$ in the following).

The parameter ϑ is to be estimated given the sample path

$$Y^T = (Y_t, 0 \leq t \leq T) \ .$$

LARGE SAMPLE MLE

IBRAGIMOV-KHASMINSKII PROGRAM

To obtain the asymptotic behavior of the implicit MLE, it is helpful to use the method by (Ibragimov & Khasminskii, 81) that consists to check several conditions on the likelihood ratio

$$\mathcal{Z}_T(u, Y^T) = \frac{\mathcal{L}_T\left(\vartheta + \frac{u}{\sqrt{T}}, Y^T\right)}{\mathcal{L}_T\left(\vartheta, Y^T\right)} \quad , \quad \vartheta + \frac{u}{\sqrt{T}} \in \mathbf{K} \subset \Theta \, .$$

Namely, we have

$$\mathcal{Z}_T(u, Z^{O,T}) \overset{law}{\Longrightarrow} \exp\left(u.\eta - \frac{u^2}{2}\mathcal{I}(\vartheta)\right) = Z(u)$$

with $\eta \sim \mathcal{N}\left(0, \mathcal{I}(\vartheta)\right)$, and for some $C, \chi > 0$, for all u such that $\vartheta + \frac{u}{\sqrt{T}} \in \mathbf{K} \subset \mathbb{R}_*^+$,

$$\mathbf{E}_\vartheta\left(\sqrt{\mathcal{Z}_T(u, Z^{O,T})}\right) \leq C \exp\left(-\chi u^2\right) \quad ,$$

and there exists $C > 0$ such that, for all $|u_1| < R$ and $|u_2| < R$,

$$\mathbf{E}_\vartheta\left(\sqrt{\mathcal{Z}_T(u_1, Z^{O,T})} - \sqrt{\mathcal{Z}_T(u_2, Z^{O,T})}\right)^2 \leq C|u_1 - u_2|^2 \, .$$

Markovian signal case

In the case where X is a Markov process and the observation process is given by

$$Y_t = Y_0 + \int_0^t h(X_s)\, ds + \sigma W_t, \quad t \geq 0 .$$

In this case, thanks to Liptser-Shyraev, the log-likelihood have the explicit form:

$$\ln \mathcal{L}_T\left(\vartheta, Y^T\right) = \int_0^T \pi_t^\vartheta\left(h(X)\right) dY_t - \frac{1}{2}\int_0^T \pi_t^\vartheta\left(h(X)\right)^2 dt$$

where $\left(\pi_t^\vartheta\left(h(X)\right), t \geq 0\right) = \left(\mathbf{E}_\vartheta\left(h(X_t)|\mathcal{F}_t^Y\right), t \geq 0\right)$ is the conditional expectation process or filtering process.

Unfortunately, we have no explicit form of the MLE due to the dependence in ϑ of the filtering process. Nevertheless, we have nonlinear nonstationnary filtering equation for $\pi_t^\vartheta\left(h(X)\right)$ which is a stochastic integral equation.

FIRST EXAMPLE: DIFFUSION LINEAR CASE

MLE large samples properties in the classical linear case

$$\begin{cases} X_t = \alpha(\vartheta) \int_0^t X_s ds + \beta V_t \\ Y_t = h \int_0^t X_s ds + \sigma W_t, \end{cases} \tag{1}$$

where $(W_t, V_t, t \geq 0)$ are two independent Wiener processes, have been treated in (Kutoyants, 80, 04).

Assuming appropriate identifiability and smoothness conditions on the coefficients, the proof relies on the following facts: the conditional expectation follows a system of stochastic differential equations, know as the Kalman-Bucy filters, and, for a negative valued function $\alpha(.)$, the filtering process have ergodic properties.

Then the MLE $\hat{\vartheta}_T$ is uniformly consistent, asymptotically normal (with regular $\sqrt{T}$-rate) with known limit variance and with converging moments.

FRACTIONAL SETTING

Unfortunately, stylised facts long-range dependance that are observed in application fields are still ignored in the previous model.

We propose to study the fractional analogue of the partial observation setting, where the observation process is given by:

$$Y_t = Y_0 + \int_0^t h(X_s)ds + \sigma W_t^H , \quad t \geq 0 ,$$

where $h(.)$ is a known function and $\sigma > 0$ a known constant, $W_t^H , t \geq 0$ is a fractional Brownian motion (fBm) valued in $\mathbb{R}$, zero-mean Gaussian process of covariance function

$$\mathbb{E} W_s^H W_t^H = \frac{1}{2} \left(t^{2H} + s^{2H} - |t-s|^{2H} \right)$$

with known Hurst exponent $H \in (0, 1)$ and independent of X, and Y_0 is a the initial condition (say $Y_0 = 0$ in the following).

When $H \neq \frac{1}{2}$ and contrary to the Wiener case, W^H is no more a Markov process nor a semimartingale and the likelihood can not be written directly in its explicit form.

FRACTIONAL LINEAR CASE

Let us consider, the linear case

$$Y_t = Y_0 + h \int_0^t X_s ds + \sigma W_t^H , \quad t \geq 0 .$$

Knowing H, even if the fBm is not a martingale, there are simple integral transformations which change the fBm to martingales (see Norros et al., 99). For some kernel $k_H(t,s)$,

$$M_t = \int_0^t k_H(t,s) dW_s^H ,$$

is a Gaussian martingale, called in the *fundamental martingale* with deterministic known variance function, and whose natural filtration of the martingale M coincides with the natural filtration of the fBm W^H.

For a signal $(X_t, t \geq 0)$ sufficiently Hölder regular, it is possible to introduce $Z = (Z_t, t \geq 0)$ the *fundamental semi-martingale* associated to Y, namely

$$Z_t = \int_0^t k_H(t,s) dY_s . \qquad (2)$$

It can be shown that the following representation holds:

$$Z_t = Z_0 + \lambda h \int_0^t Q_s d\langle M\rangle_s + \sigma M_t \,,$$

where

$$Q_t = \frac{d}{d\langle M\rangle_t} \int_0^t k_H(t,s) X_s ds \,.$$

Therefore, thanks to Lipster-Shyraev, the log-likelihood can be written

$$\ln \mathcal{L}_T\left(\vartheta, Z^T\right) = \lambda h \int_0^T \pi_t^{\vartheta}(Q)\, dZ_t - \frac{(\lambda h)^2}{2} \int_0^T \pi_t^{\vartheta}(Q)^2\, d\langle M\rangle_t$$

where $\pi_t^{\vartheta}(Q) = \mathbf{E}_{\vartheta}\left(Q_t | \mathcal{F}_t^Z\right)$ is the conditional expectation and $\mathcal{F}_t^Z$ the natural filtration, $\mathcal{F}_t^Z = \sigma\left(Z_s, 0 \le s \le t\right)$.

FIRST EXAMPLE

MLE properties in the fractional linear case

$$
\begin{cases}
X_t = X_0 - \vartheta \int_0^t X_s ds + \beta V_t^H \\
Y_t = Y_0 + h \int_0^t X_s ds + \sigma W_t^H,
\end{cases}
\tag{3}
$$

where $\vartheta > 0$ is unknown and $W_t^H, V_t^H, t \geq 0$ are two independent fBm with **same** known Hurst parameter, have been treated in (B. and Kleptsyna, 08) and with dependent noise in (B., 09).

The proof rely on the fact that $(Q_t, t \geq 0)$ can be written as $(\ell(t)\zeta_t, t \geq 0)$ where $(\zeta_t, t \geq 0)$ is a two-dimensional process whose conditional expectation $\pi_t^{\vartheta}(\zeta)$ has an explicit system of stochastic differential equations known as the *Kalman-Bucy filters* shown by (Kleptsyna and Le Breton, 02).

Actually, ζ is not ergodic, but, when $\vartheta > 0$, Q has ergodic properties. The knowledge of likelihood quadratic term behaviour via Laplace transform allows us to check all the properties of the likelihood ratio required in the Ibragimov-Khasminskii program and to deduce the large samples MLE properties.

SECOND EXAMPLE

The signal is deterministic but controlled

$$\begin{cases} X_t = X_0 + \int_0^t (-\vartheta X_s + u_s)ds \\ Y_t = Y_0 + h\int_0^t X_s ds + \sigma W_t^H, \end{cases} \tag{4}$$

In the classical case ($H = \frac{1}{2}$), the estimation problem of the drift coefficient have been treated in (Ovseevich, Khasminskii and Chow, 00).

The key point of this work is to find an optimal control u such that

$$\mathcal{J}_T(\vartheta) = \mathcal{I}_T(\vartheta, u^*) = \sup_{u\in\mathcal{U}_T} \mathcal{I}_T(\vartheta, u) ,$$

where $\mathcal{I}_T(\vartheta, u)$ stands for the Fischer information and deduced an estimator $\overline{\vartheta}_T$ (based on MLE) of the parameter ϑ which are asymptotically efficient in the sense that, for any compact $\mathbb{K} \subset \mathbb{R}^+$,

$$\sup_{\vartheta\in} \mathcal{J}_T(\vartheta)\mathbf{E}_\vartheta \left(\overline{\vartheta}_T - \vartheta\right)^2 = 1 + o(1) , \tag{5}$$

as $T \to \infty$.

In the fractional case ($H \neq \frac{1}{2}$), after the transformation, the problem of finding an optimal input is the following.
The Fischer information stands for

$$\begin{aligned} \mathcal{J}_T(\vartheta) &= \sup_{v \in \mathcal{V}_T} \mathcal{I}_T(\vartheta, v) \\ &= T \sup_{\tilde{v} \in L^2([0,T]),\, ||\tilde{v}|| \leq 1} \int_0^T \int_0^T R_T(s,\sigma)\tilde{v}(s)\tilde{v}(\sigma) ds d\sigma. \end{aligned}$$

PROPOSITION

$$\lim_{T \to +\infty} \sup_{\tilde{v} \in L^2([0,T]),\, ||\tilde{v}|| \leq 1} (R_T\tilde{v}, \tilde{v}) = \frac{\mu^2}{\vartheta^4}.$$

Moreover the asymptotical optimal input in the class of controls $\mathcal{U}_T$ is

$$u_{opt}(t) = \frac{\kappa_H}{\sqrt{2\lambda}} t^{H-\frac{1}{2}}$$

where the constants λ and κ_H are defined .

- To prove the lower bound it is sufficient to take

$$\tilde{v}_t = \frac{\mathbf{1}(t \in [0, T])}{\sqrt{T}}.$$

- To prove the upper bound we have to show that the first eigenvalue $\nu_1(T)$ of the compact self-adjoint operator

$$\begin{aligned} \mathcal{K}_T : \ L^2([0,T]) &\longrightarrow L^2([0,T]) \\ \tilde{v} &\longmapsto \mathcal{K}_T \tilde{v} = \int_0^T R_T(t,s)\tilde{v}(s)ds \end{aligned}$$

is bounded by $\frac{\mu^2}{\vartheta^4}$, for T sufficiently large.

The estimation of the first eigenvalue $\nu_1(T)$ is based on the special structure of the kernel $R_T(s,\sigma)$. We introduce the centered Gaussian process $\xi = \left(\xi_t^1, \xi_t^2\right), 0 \le t \le T$ with a second component $\xi^2$with a covariance function closely related to $R_T(s,\sigma)$. Then ξ is the solution of the following backward stochastic differential equation. Now we can use the general representation of the Laplace transform

$$\begin{aligned} L_T(a) &= \mathbf{E}_\vartheta \exp\left\{ -a \int_0^T \left[\frac{\mu}{2} \xi_t^2 b(t) t^{\frac{1}{2}-H} \right]^2 dt \right\} \\ &= \prod_{i \ge 1} (1 + 2a\nu_i(T))^{-\frac{1}{2}}, \end{aligned}$$

where $(\nu_i, i \ge 1)$ - the eigenvalues of the operator $\mathcal{K}_T$.

- A. Brouste & M. Kleptsyna (2010) *Asymptotic properties of MLE for partially observed fractional diffusion system*, Statistical Inference for Stochastic Processes, 13(1), 1–13.
- A. Brouste (2010) *Asymptotic properties of MLE for partially observed fractional diffusion system with*