Nonparametric estimation problem for a periodic stochastic process

Khalil EL WALED
Université Rennes 2

Abstract

In this paper we construct a kernel estimator of a periodic signal when the observation follows the model $d\zeta_t = f(t)dt + \sigma(t)dW_t$, where $f, \sigma : \mathbb{R} \to \mathbb{R}$ are continuous periodic and $\{W_t, t \geq 0\}$ is a Brownian motion. We state its consistency as well as the asymptotic normality.

1 Introduction

We consider the following model of periodic signal disturbed by noise

$$d\zeta_t = f(t)dt + \sigma(t)dW_t, \qquad t \geq 0, \tag{1}$$

where $f, \sigma : \mathbb{R} \mapsto \mathbb{R}$ are two continuous periodic functions with the same period P, and $W = \{W_t, t \geq 0\}$ is a Brownian motion. Here we focus on the estimation of the time periodic drift $f(\cdot)$ when we observe a realization of process (1) along a time interval $[0, T]$ as $T \to \infty$. We will see that this statistical problem is an i.i.d. nonparametric estimation problem.
As an application, let $\{\xi_t, t \geq 0\}$ be the time dependent geometric Brownian motion which verifies the following linear stochastic differential equation

$$d\xi_t = f(t)\xi_t dt + \sigma(t)\xi_t dW_t. \tag{2}$$

Then the observation of $\{\xi_t, t \in [0, T]\}$ is equivalent to the observation of $\{\zeta_t, t \in [0, T]\}$, and the estimation of $f(\cdot)$ is identical in model (2) and in model (1).
Equations of such a type arise in many domains for instance in finance (Karatzas and Shreve, 1991) (Black-Scholes-Merton model), mechanics (Has'minskiĭ, 1980) and in biology (Collet and Martinez, 2008).

2 Properties of $\{\zeta_t, t \geq 0\}$

Process $\{\zeta_t, t \geq 0\}$ is a Gaussian process with independent increments, $F(t) := E[\zeta_t] = \int_0^t f(u)\,du$, $var[\zeta_t] = \int_0^t \sigma^2(u)\,du$. Moreover we have

$$\zeta_{nP+t} = nF(P) + F(t) + \sum_{k=0}^{n-1} Z_k + \mathbf{Z}_n(t) \tag{3}$$

for all $n \in \mathbb{N}$ and $t \in [0, P]$, where $\mathbf{Z}_k(t) := \int_0^t \sigma(u)\,dW_u^{(kP)}$, $Z_k := \mathbf{Z}_k(P)$ and $W_u^{(kP)} := W_{kP+u} - W_{kP}$.
The processes $\{W_u^{(kP)}, u \in [0, P]\}$, $k \geq 0$ are Brownian motions and the processes $\mathbf{Z}_k := \{\mathbf{Z}_k(u), u \in [0, P]\}$, $k \geq 0$, are i.i.d. in $\mathcal{C}[0, P]$. Thus the process $\{\zeta_t, t \geq 0\}$ is an inhomogeneous Markov process with a periodic transition semigroup. Following (Höpfner and Kutoyants, 2010) define the P-segments time series

$$\mathbf{Y}_n := \{\zeta_{nP+t}, t \in [0, P]\}, \qquad n \in \mathbb{N}$$

which fulfills a functional autoregressive representation

$$\mathbf{Y}_n = \mathbf{Y}_{n-1}(P) + F(\cdot) + \mathbf{Z}_n.$$

So $(\mathbf{Y}_n)_{n \in \mathbb{N}}$ is a homogeneous Markov sequence with state space $\mathcal{C}[0, P]$. As $Z_k = \mathbf{Z}_k(P), k \in \mathbb{N}$ are i.i.d., the (SLLN) applies

$$\lim_{n\to\infty} \sup_{t\in[0,P]} \left| \frac{1}{n}\mathbf{Y}_n(t) - F(P) \right| = 0 \quad \mathrm{P-a.e.}$$

Then $\mathrm{P-a.e.}$ limits

$$\lim_{t\to\infty} \zeta_t = \begin{cases} -\infty & \text{if } F(P) < 0 \\ \infty & \text{if } F(P) > 0. \end{cases}$$

When $F(P) = 0$, we obtain (LiL) that , for $\zeta_n^P = \frac{\zeta_{nP}}{\sqrt{2n \ln \ln n}}$

$$\mathrm{P}\left[\liminf_{n\to\infty} \zeta_n^P = -\sqrt{G(P)}\right] = \mathrm{P}\left[\limsup_{n\to\infty} \zeta_n^P = \sqrt{G(P)}\right] = 1$$

where $G(P) := \int_0^P \sigma^2(u)\,du$. As $\sigma(\cdot)$ is continuous but not identically null, $G(P) > 0$ and

$$\mathrm{P}\left[\liminf_{t\to\infty} \zeta_t = -\infty\right] = \mathrm{P}\left[\limsup_{t\to\infty} \zeta_t = \infty\right] = 1.$$

Thus the Markov process $\{\zeta_t : t \geq 0\}$ is recurrent when $F(P) = 0$, and transient otherwise (see Has'minskiĭ, 1980).

3 Estimation of $f(\cdot)$

We assume that we observe one realization of the process (1) along $[0, T]$, as $T \to \infty$. $\sigma(\cdot)$, P are known. The target is to estimate $f(\cdot)$.

Periodic kernel

Consider $K : \mathbb{R} \to \mathbb{R}$ continuous, symmetric with compact support contained in $[-\frac{1}{2}, \frac{1}{2}]$. Assume that

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} K(u)du = 1.$$

Let $h > 0$ and define the function

$$K_h(t) := \frac{1}{h} \sum_{k\in\mathbb{Z}} K\left(\frac{t + kP}{h}\right).$$

Lemma 1. *For each $h \in (0, P)$, $K_h(\cdot)$ is periodic with period P.*

1. $supp(K_h) \subset \bigcup_{k\in\mathbb{Z}} \left[kP - \frac{h}{2}, kP + \frac{h}{2}\right]$;

2. $\int_{-\frac{h}{2}}^{\frac{h}{2}} K_h(u)du = 1$.

Definition of the estimator

The kernel estimator of $f(t)$ is defined by

$$\hat{f}_n(t) := \frac{1}{n} \int_0^{nP} K_{h_n}(t - u)\,d\zeta_u$$

where $P > 0$ is the known period of $f(\cdot)$, $(h_n)_{n\in\mathbb{N}} \searrow 0$. The estimator kernel $K_{h_n}(\cdot)$ is periodic

$$\begin{aligned} \hat{f}_n(t) &= \frac{1}{n}\int_0^{nP} K_{h_n}(t-u)f(u)du + \frac{1}{n}\int_0^{nP} K_{h_n}(t-u)\sigma(u)dW_u \\ &:= U_n(t) + V_n(t). \end{aligned}$$

Thus $\mathrm{E}\left[\hat{f}_n(t)\right] = U_n(t)$ and the estimator is asymptotically unbiased.

Lemma 2. *With the above notations we have*

1. $U_n(\cdot)$ is a continuous periodic function with period P;

2. $\lim_{n\to\infty} \sup_{t\in[0,P]} |U_n(t) - f(t)| = 0$.

The continuity and the periodicity of the function $U_n(\cdot)$ are direct consequences from $K_{h_n}(\cdot)$ ones. Moreover

$$U_n(t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} K(v) f(t - vh_n)\,dv.$$

As $f(\cdot)$ is periodic and continuous,

$$\sup_{t\in[0,P]} |U_n(t) - f(t)| \leq \int_{-\frac{1}{2}}^{\frac{1}{2}} K(v) \sup_{t\in[0,P]} |f(t - vh_n) - f(t)|\,dv$$

Remark Furthermore, following Tsybakov (2004, chapter 1) we assume that $f(\cdot)$ is in Hölder class $\mathcal{H}^\beta$ for some smoothness order $\beta \geq 1$, and we consider a kernel $K(\cdot)$ of order $l := \lfloor \beta \rfloor$, the biggest integer which is strictly smaller than the real β, i.e. $\int u^j K(u)du = 0, j = 1, \cdots, l$. Then we can control Taylor expansion of order l of $f(\cdot)$ and we deduce

$$|U_n(t) - f(t)| \leq Ch_n^\beta \tag{4}$$

where

$$C = \frac{L}{l!} \int_{-\frac{1}{2}}^{\frac{1}{2}} |v|^\beta |K(v)| dv$$

and L is a positive constant depending on $f(\cdot)$. This gives an order $O(h_n^\beta)$ for the bias terms. Moreover we get

$$\limsup_{n\to\infty} h_n^{-\beta} \sup_{t\in[0,P]} |U_n(t) - f(t)| < \infty.$$

In particular for $h_n = o\left(n^{-\frac{1}{2\beta+1}}\right)$

$$\lim_{n\to\infty} \sqrt{nh_n} \sup_{t\in[0,P]} |U_n(t) - f(t)| = 0. \tag{5}$$

4 Consistency, Asymptotic normality

Mean square convergence

Under the assumption that $f(\cdot)$ is continuous, the estimator $\hat{f}_n(t)$ converges in mean square to $f(t)$. Moreover we have

$$\lim_{n\to\infty} \sup_{t\in[0,P]} \mathrm{E}\left[\left|\hat{f}_n(t) - f(t)\right|^2\right] = 0, \tag{6}$$

$$\mathrm{E}\left[\left|\hat{f}_n(t) - f(t)\right|^2\right] \leq C^2 h_n^{2\beta} + \frac{1}{nh_n} \sup_u \sigma^2(u) \int_{-\frac{1}{2}}^{\frac{1}{2}} K^2(v)\,dv.$$

Theorem 1. *Assume that $f(\cdot) \in \mathcal{H}^\beta$ and $K(\cdot)$ is of order $l := \lfloor \beta \rfloor$ for some $\beta > 1$. Then if $h_n = o(n^{-\frac{1}{2\beta+1}})$, we have*

$$\lim_{n\to\infty} \sup_{t\in[0,P]} \left| nh_n \mathrm{E}\left[\left|\hat{f}_n(t) - f(t)\right|^2\right] - \sigma^2(t) \int_{-\frac{1}{2}}^{\frac{1}{2}} K^2(v)dv \right| = 0.$$

Hence, we obtain the rate of the (MISE) convergence

$$\lim_{n\to\infty} nh_n \int_0^P \mathrm{E}\left[\left|\hat{f}_n(t) - f(t)\right|^2\right] dt = \left(\int_0^P \sigma^2(t)\,dt\right)\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} K^2(v)\,dv\right).$$

Notice that as for the smoothness of the diffusion function $\sigma(\cdot)$, we need only that this function is continuous.

Asymptotic normality

Using the fact that the process $\bar{f}_n(t) := \sqrt{nh_n}\left(\hat{f}_n(t) - f(t)\right)$ is a Gaussian process we obtain

Theorem 2. *Assume that the hypotheses of Theorem 1 are satisfied. Then for any integer $k \geq 1$ and all $0 \leq t_1 < \cdots < t_k < P$, the random vector $\left(\bar{f}_n(t_1), \ldots, \bar{f}_n(t_k)\right)$ is asymptotically normal:*

$$\lim_{n\to\infty} \mathcal{L}\left(\bar{f}_n(t_1), \ldots, \bar{f}_n(t_k)\right) = \mathcal{N}_k\left(0, \Sigma_{t_1,\ldots,t_k}\right)$$

where $\mathcal{N}_k\left(0, \Sigma_{t_1,\ldots,t_k}\right)$ is the zero-mean k-dimensional normal law the variance matrix of which is equal to the $k \times k$-diagonal matrix

$$\Sigma_{t_1,\ldots,t_k} = \int_{-1}^{1} K^2(u)du \times \mathrm{diag}\left(\sigma^2(t_1), \ldots, \sigma^2(t_k)\right).$$

Strong consistency

Using Borel Cantelli lemma and martingales proprties we show the strong consistency.

Theorem 3. *Assume that $K(u) = 2(1 - 2|u|)$, for $u \in [-\frac{1}{2}, \frac{1}{2}]$, $K(u) = 0$ otherwise. Let $h_n = n^{-a}$, $0 < a < 1$, Then for each t,*

$$\lim_{n\to\infty} \hat{f}_n(t) = f(t) \qquad \mathrm{P-}a.e.$$

5 Simulation

We take the particular values $P = 1$, $T = nP = 1000$, $h_n = 10^{-2}$. Boxplots of estimations of $f(\cdot)$ from 20 simulations.

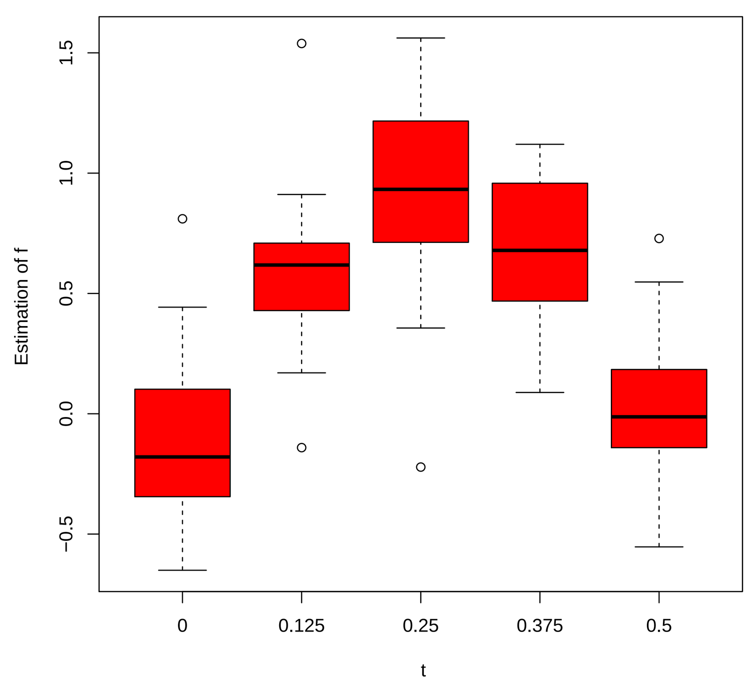

Figure 1: $d\zeta_t = sin(2\pi t)dt + dW_t$.

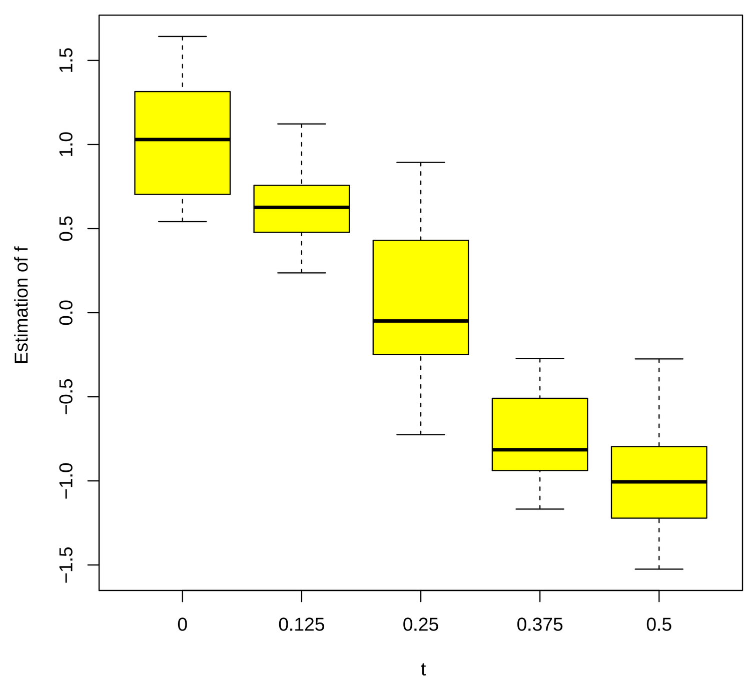

Figure 2: $d\xi_t = cos(2\pi t)\xi_t dt + \xi_t dW_t$.

This poster is taken from the paper: D. Dehay, K. El Waled. Statist. Probab. Letters 83 (2013) 608-615.

References

[1] Collet, P., Martinez, S., 2008. Asymptotic velocity of one dimensional diffusions with periodic drift, *J. Math. Biol.* 56 765–792.

[2] Has'minskiĭ, R.Z., 1980. *Stochastic Stability of Differential Equations*, Sijthoff & Noordhoff, Alphaen aan den Rijn.

[3] Höpfner, R., Kutoyants, Yu., 2010. Estimating discontinuous periodic signals in a time inhomogeneous diffusion. *Statist. Inference Stoch. Process.* 13 193–230.

[4] Karatzas, I., Shreve, S.E., 1991. *Brownian Motion and Stochastic Calculus*, 2nd edition, Springer-Verlag, New York.

[5] Tsybakov, A.B., 2004. *Introduction à l'estimation Non-Paramétrique* , Springer, Berlin.