Parameter estimation for stationary solutions of stochastic delay equations

Alexander Gushchin

(based on a joint work with Uwe Küchler)

Le Mans, March 23, 2007

Let $J := [-r, 0]$, where $r \in [0, \infty)$, and

$$a \in \mathbb{M} := \{\text{all finite signed measures on } J\}.$$

Consider the equation

$$\begin{aligned} dX(t) &= \int_J X(t+u)a(du)dt + dW(t), && t \geq 0, \\ X(t) &= X_0(t), && t \in J, \end{aligned} \tag{1}$$

where W is a standard Brownian motion, X_0 is a continuous process on J *independent* of W. The equation has a unique pathwise continuous solution with the distribution determined by the delay measure a and the distribution of X_0. Put

$$h_a(\lambda) := \lambda - \int_J e^{\lambda u} a(du), \quad \lambda \in \mathbb{C}.$$

The equation (1) has a *stationary* solution if and only if $a \in \mathbb{M}^-$, where

$$\mathbb{M}^- := \{a \in \mathbb{M} \colon h_a(\lambda) \neq 0 \text{ for } \operatorname{Re}\lambda \geq 0\}$$

(Gushchin and Küchler, 2000).

The stationary solution is a Gaussian process with the spectral density

$$f_a(\lambda) = \frac{1}{2\pi |h_a(i\lambda)|^2}$$

and the covariance function

$$K_a(t) = \int_{\mathbb{R}} e^{i\lambda t} f_a(\lambda)\, d\lambda.$$

If $a = -\vartheta\delta_{\{0\}}$, $\vartheta > 0$, then $a \in \mathbb{M}^-$ and

$$f_a(\lambda) = \frac{1}{2\pi(\vartheta^2 + \lambda^2)}, \qquad K_a(t) = \frac{1}{2\vartheta} e^{-\vartheta|t|}.$$

For an arbitrary $a \in \mathbb{M}^-$ the functions f_a and K_a have similar properties:

$$\frac{B_*}{1+\lambda^2} \le f_a(\lambda) \le \frac{B^*}{1+\lambda^2}, \qquad |K_a(t)| \le Le^{-\gamma t}, \tag{2}$$

$0 < B_* < B^*$, $\gamma > 0$, $L > 0$,

$$f_a(\lambda) \sim \frac{1}{2\pi(1+\lambda^2)} \quad \text{as} \quad \lambda \to \pm\infty, \tag{3}$$

K_a is differentiable everywhere except 0, $K_a'(\mp 0) = \pm 1/2$, and the derivative K_a' satisfies the Lipschitz condition on $(0, \infty)$ and $(-\infty, 0)$.

For $a \in \mathbb{M}^-$ let P^a be the distribution of the stationary solution to the equation (1) in the space $\mathbf{C}[-r,\infty)$ of continuous functions on $[-r,\infty)$, and let P^a_T be its restriction to the σ-field corresponding to observations on $[-r,T]$.

Proposition 1 *Let $a, a' \in \mathbb{M}^-$ and $b := a' - a$. Then $P^a_T \sim P^{a'}_T$ for any $T \in \mathbb{R}_+$ and*

$$\frac{dP^{a'}_T}{dP^a_T}(X) = \frac{dP^{a'}_0}{dP^a_0}(X) \exp\left\{ \int_0^T Y(t)\, dW(t) - \frac{1}{2} \int_0^T Y^2(t)\, dt \right\},$$

where X is the stationary solution to the equation (1) and

$$Y(t) := \int_J X(t+u)\, b(du).$$

In what follows all delay measures a (with different indices) are assumed to belong to a subset $\mathcal{A}$ of $\mathbb{M}^-$ such that

$$\sup_{a\in\mathcal{A}} \|a\|_{TV} < \infty$$

and the closure of $\mathcal{A}$ with respect to the dual-Lipschitz norm is still in $\mathbb{M}^-$; here

$$\|b\|_{DL} = \sup \left| \int_J f(u)\, b(du) \right|, \quad b \in \mathbb{M},$$

where the supremum is taken over all functions f on J satisfying $|f| \leq 1$ and $|f(u) - f(u')| \leq |u - u'|$. The relations (2) and (3) are uniform over $\mathcal{A}$.

Statistical problem

Assume that we observe a continuous realization of a stationary solution to the affine stochastic delay differential equation

$$dX(t) = \int_{[-r,0]} X(t+u) a_\vartheta(du)\, dt + dW(t), \quad t \geq 0,$$

with the initial condition $X(t) = X_0(t)$, $t \in [-r, 0]$, independent of the standard Wiener process W; here $\vartheta \in \Theta$, where Θ is an open subset of $\mathbb{R}^k$, and $a_\vartheta \in \mathcal{A}$, $\vartheta \in \Theta$. We are interested in asymptotic inference about the parameter ϑ as the duration of observations $T \to \infty$.

In the next theorem

- $\vartheta_0 \in \Theta$;

- (φ_T) is a (normalizing) family of non-singular matrices of the form

$$\varphi_T = \varphi^{(1)} \varphi_T^{(2)},$$

where

$$\varphi_T^{(2)} = \mathrm{diag}[\delta_{T,1}, \ldots, \delta_{T,k}],$$

$$\delta_{T,i} \downarrow 0, \quad T \to \infty, \quad \forall i = 1, \ldots, k;$$

- $$Z_T(u) = \log \frac{dP_T^{a_{\vartheta_0 + \varphi_T u}}}{dP_T^{a_{\vartheta_0}}}, \quad u \in \mathbb{R}^k.$$

Theorem 1 *Assume that the finite-dimensional distributions of* $Z_T(u)$ *under* $P_T^{a_{\vartheta_0}}$ *weakly converge (uniformly over compacts) to the finite-dimensional distributions of a limiting random function* $Z(u)$ *as* $T \to \infty$, *where* $Z(u)$ *is continuous in probability and* $EZ^2(u) > 0$ *for all* $u \neq 0$. *Then* $Z(u) = B(u) - \frac{1}{2}EB^2(u)$, *where* $B(u)$ *is a centered Gaussian function, there exist positive numbers* $p_1, \ldots, p_k$ *such that*

$$B(\lambda^{p_1}u_1, \ldots, \lambda^{p_k}u_k) \overset{law}{=} \lambda B(u_1, \ldots, u_k) \quad \forall \lambda > 0,$$

and $\delta_{T,i} = T^{-p_i/2}\ell_i(T)$, *where* $\ell_i(T)$ *are slowly varying functions as* $T \to \infty$, *i.e.* $\ell_i(\mu T)/\ell_i(T) \to 1$ *for any* $\mu > 0$, $i = 1, \ldots, k$.

The case $p_1 = \cdots = p_k = 1$ and $B(u) = \langle u, \xi \rangle$, $\xi \sim N(0, I)$, corresponds to the local asymptotic normality (LAN). It was shown in (Gushchin and Küchler, 2003) that this is the case with $\delta_{T,1} = \cdots = \delta_{T,k} = T^{-1/2}$ if the mapping $\vartheta \rightsquigarrow a_\vartheta$ is smooth enough (roughly speaking, it is differentiable in a certain sense with a non-singular gradient).

Küchler and Kutoyants (2000) considered the following model:

$$a_\vartheta = b\delta_{\{\vartheta\}}, \quad b < 0 \text{ is known}, \quad \frac{\pi}{2b} < \vartheta < 0.$$

Then $p = 2$, $\delta_T = T^{-1}$, $B(u) = W(u)$, where $W(u)$ is the (two-sided) Brownian motion.

Kutoyants (2004) considered the same model, where now b is also unknown. Then $p_1 = 1$, $p_2 = 2$, $\delta_{T,1} = T^{-1/2}$, $\delta_{T,2} = T^{-1}$, $B(u_1, u_2) = u_1\xi + W(u_2)$, where $\xi \sim N(0,1)$ is independent of $W(u)$.

The model

Let b be a finite signed measure on $\mathbb{R}$ with a compact support. Let b_ϑ be the image of b under the shift $x \rightsquigarrow x + \vartheta$, i.e. b_ϑ is the translate of b by ϑ. Then, if r is large enough, $\Theta' := \{\vartheta \colon b_\vartheta \in \mathbb{M}\}$ contains an open interval. Without loss of generality assume that 0 is in this interval.

However, the set $\{\vartheta \colon b_\vartheta \in \mathbb{M}^-\}$ can be empty.

Now let $a \in \mathbb{M}^-$. Put

$$a_\vartheta := a + b_\vartheta - b.$$

Let Θ be an open bounded interval containing 0 such that $a_\vartheta \in \mathbb{M}^-$ for $\vartheta \in \overline{\Theta}$. Such an interval always exists.

If $a = b$, then $a_\vartheta = b_\vartheta$, and we have the problem of estimating *delay location*.

Let

$$\varphi_b(\lambda) = \int_J e^{i\lambda u}\, b(du), \qquad \lambda \in \mathbb{R},$$

be the Fourier transform of b. Put

$$H := \sup\left\{\gamma \le 1 \colon \int_{-\infty}^{\infty} (1+\lambda^2)^{-1} |\lambda|^{2\gamma} |\varphi_b(\lambda)|^2 d\lambda < \infty\right\}.$$

It is clear that $1/2 \le H \le 1$. It turns out that if the assumptions of Theorem 1 are satisfied, then necessarily $p = H^{-1}$. However, it may happen that there is no proper normalization φ_T such that the assumptions of Theorem 1 are satisfied.

Put

$$\Psi_b(x) := \int_{-x}^{x} |\varphi_b(\lambda)|^2 d\lambda, \quad x \in [0, \infty),$$

and

$$\delta_T^{-1} = \inf\{x \geq 1 \colon Tx^{-2}\Psi_b(x) < 1\}.$$

Lemma 1 *Let $\Psi_b(x)$ be regularly varying (at infinity) of index β, i.e. $\Psi_b(x) = x^\beta L(x)$, where $L(x)$ is slowly varying. Then*

$$\beta = 2 - 2H$$

and δ_T is monotone decreasing to 0 and varies regularly with index $-\frac{1}{2H}$; moreover,

$$\delta_T = O(T^{-1/2}) \quad \text{if } H = 1 \qquad \text{and} \qquad T^{-1} = O(\delta_T) \quad \text{if } H = 1/2.$$

From now on we shall write P^{ϑ} instead of P^{a_ϑ}. Recall that

$$Z_T(u) = \log \frac{dP_T^{\vartheta_0+\varphi_T u}}{dP_T^{\vartheta_0}}, \quad u \in \mathbb{R}.$$

Proposition 2 *Assume that for some $\vartheta_0 \in \Theta$ there is a normalizing function $\varphi_T \to 0$, $T \to \infty$, such that the distribution of* $\exp(Z_T(u))$ *under P^{ϑ_0} weakly converges to the distribution of $V(u)$ for every $u \in \mathbb{R}$, where $V(u) \not\equiv 1$ and $V(u)$ tends to 1 in distribution as $u \to 0$. Then $\Psi_b(x)$ is regularly varying*, *and there is a constant $c \in (0, \infty)$ such that $\varphi_T \sim c\delta_T$, $T \to \infty$.*

Theorem 2 *Assume that* $\Psi_b(x)$ *is a regularly varying function as* $x \to \infty$ *and* $\varphi_T = \delta_T$. *Then uniformly in* $\vartheta_0 \in \Theta$ *the finite-dimensional distributions of the random function* $Z_T(u)$ *under* P^{ϑ_0} *weakly converge (uniformly in* u *over compact intervals) to the finite-dimensional distributions of the random function* $Z(u)$, *where*

$$Z(u) := cB^H(u) - \frac{c^2}{2}E[B^H(u)]^2$$

and $B^H(u)$ *ia a standard two-sided fractional Brownian motion with Hurst index* H.

Remarks: 1. The constant c in Theorem 2 can be written explicitly. Itdepends only on H except the (“regular”) case where $\delta_T \asymp T^{-1/2}$ or, equivalently, $\Psi_b(x)$ is bounded. In the latter case c may depend on a, b and ϑ_0.

2. $H = 1$ and $\delta_T \asymp T^{-1/2}$ if and only if b is absolutely continuous with respect to the Lebesgue measure with the square-integrable density. This improves the result in Gushchin and Küchler (2003) for this particular model.

3. $H = 1/2$ and $\delta_T \asymp T^{-1}$ if and only if b has at least one atom. This generalizes the result in Küchler and Kutoyants (2000).

Theorem 3 *Let the assumptions of Theorem 2 be satisfied and let $\widehat{\vartheta}_T$ be the maximum likelihood estimator. Then uniformly in* $\vartheta_0 \in \Theta$

(1)

$$\mathrm{Law}(\delta_T^{-1}(\widehat{\vartheta}_T - \vartheta_0) \mid P^{\vartheta_0}) \Rightarrow \mathrm{Law}(\widehat{u}), \quad T \to \infty,$$

where $\widehat{u}$ is a point of maximum of the limit function $Z(u)$;

(2) *all the moments of $\delta_T^{-1}(\widehat{\vartheta}_T - \vartheta_0)$ under P^{ϑ_0} converge as $T \to \infty$ to the corresponding moments of $\widehat{u}$;*

(3) *for $\gamma \in (0, 2H)$ there are positive constants b_0 and B_0 such that for T large enough*

$$\sup_{\vartheta_0 \in \Theta} P^{\vartheta_0}\{\delta_T^{-1}|\widehat{\vartheta}_T - \vartheta_0| > R\} \leq B_0 \exp\left(-b_0 R^\gamma\right), \quad R > 0.$$

Theorem 4 *Let the assumptions of Theorem 2 be satisfied and let $\widetilde{\vartheta}_T$ be the Bayes estimator corresponding to a continuous prior positive density $q(u)$ on $\overline{\Theta}$ and the loss function $l(u) = |u|^p$, $p \geq 1$. Then uniformly in $\vartheta_0 \in \Theta$*

$$(1) \qquad \mathrm{Law}(\delta_T^{-1}(\widetilde{\vartheta}_T - \vartheta_0) \mid P^{\vartheta_0}) \Rightarrow \mathrm{Law}(\widetilde{u}), \quad T \to \infty,$$

where $\widetilde{u}$ is the point of minimum of the function $\int_{\mathbb{R}} l(u-v)\exp(Z(v))\,dv$;

(2) all the moments of $\delta_T^{-1}(\widetilde{\vartheta}_T - \vartheta_0)$ under P^{ϑ_0} converge as $T \to \infty$ to the corresponding moments of $\widetilde{u}$;
(3) for $\gamma \in (0, 2H)$ there are positive constants b_0 and B_0 such that for T large enough

$$\sup_{\vartheta_0 \in \Theta} P^{\vartheta_0}\{\delta_T^{-1}|\widetilde{\vartheta}_T - \vartheta_0| > R\} \leq B_0 \exp\left(-b_0 R^\gamma\right), \quad R > 0;$$

(4) the family of estimators $\{\widetilde{\vartheta}_T\}$ is asymptotically efficient with respect to the class of loss functions $l(\delta_T^{-1}u)$.

Remark: If the function $\Psi_b(x)$ is not regularly varying as $x \to \infty$, we can show that for any $\gamma < 1/2$ and $\vartheta \in \Theta$

$$T^{-\gamma}(\widehat{\vartheta}_T - \vartheta) \xrightarrow{P^\vartheta} 0 \quad \text{and} \quad T^{-\gamma}(\widetilde{\vartheta}_T - \vartheta) \xrightarrow{P^\vartheta} 0.$$

Let ${}^{-}b$ be the reflection of b around zero and ${}^{0}b := b \star {}^{-}b$ its symmetrization. Put $F(x) := {}^{0}b([0, x])$, $x > 0$.

Theorem 5 *Assume that b is a positive (or a negative) measure. Then $\Psi_b(x)$ is a regularly varying function as $x \to \infty$ with the index $2-2H$, $H \in [1/2, 1]$ if and only if $F(x)$ is a regularly varying function as $x \downarrow 0$ with the index $2H - 1$.*

Theorem 6 *Assume that b is a positive measure on an interval (U, V) with an increasing (or decreasing) density p with respect to the Lebesgue measure. Then $F(x)$ is a regularly varying function as $x \downarrow 0$ with the index $2H - 1$, $H \in (1/2, 1)$ if and only if $p(x)$ is a regularly varying function with the index $H - 3/2$ at V (resp. at U).*

References

Prakasa Rao, B. L. S. (1968) Estimation of the location of the cusp of a continuous density. *Ann. Math. Statist.*, **39**, 76–87.

Ibragimov, I. A. and Has'minskii, R. Z. (1981) *Statistical estimation. Asymptotic theory*. Springer-Verlag, New York etc.

Strasser, H. (1985) Scale invariance of statistical experiments. *Probab. Math. Statist.*, **5**, 1–20.

Pflug, G. (1982) A statistically important Gaussian process. *Stochastic Processes Appl.*, **13**, 45–57.

Gushchin, A. A. and Küchler, U. (2000) On stationary solutions of delay differential equations driven by a Lévy process. *Stochastic Processes Appl.*, **88**, 195–211.

Küchler, U. and Kutoyants, Yu. A. (2000) Delay estimation for some stationary diffusion-type processes. *Scand. J. Stat.*, **27**, 405–414.

Gushchin, A. A. and Küchler, U. (2003) On parametric statistical models for stationary solutions of affine stochastic delay differential equations. *Math. Methods Statist.*, **12**, 31–61.

Kutoyants, Y. A. (2004) Statistical inference for ergodic diffusion processes, Springer, London.

Gushchin, A. A. and Küchler, U. (2004) On recovery of a measure from its symmetrization. *Theory Probab. Appl.*, **49** (2004), 352–362.

Gushchin, A. A. and Küchler, U. (2007) On estimation of delay location (in preparation).