On Multi-step MLE-processes in the Problems of Estimation of the Solution of BSDEs.

Small noise, unknown volatility and ergodic diffusion cases.

Yury A. Kutoyants

Université du Maine, Le Mans, FRANCE

National Research University "MPEI", Moscow , RUSSIA

School on Stochastics and Financial Mathematics,

Sochi, September 10, 2015

Backward Stochastic Differential Equation

Problem: We are given a stochastic differential equation (called *forward*)

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + a(t, X_t)\,\mathrm{d}W_t, \quad X_0 = x_0,\ 0 \le t \le T,$$

and two functions $f(t, x, y, z)$ and $\Phi(x)$. We have to construct a couple of processes (Y_t, Z_t) such that the solution of the equation

$$\mathrm{d}Y_t = -f(t, X_t, Y_t, Z_t)\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, \quad Y_0,\ 0 \le t \le T,$$

(called *backward*) has the final value $Y_T = \Phi(X_T)$.

For the existence and uniqueness of the solution see Pardoux and Peng (1990). The *Markovian case* considered here was discussed by Pardoux and Peng (1992) and El Karoui & al. (1997).

Solution: Suppose that $u(t,x)$ satisfies the equation

$$\frac{\partial u}{\partial t} + b(t,x)\frac{\partial u}{\partial x} + \frac{1}{2}a(t,x)^2\frac{\partial^2 u}{\partial x^2} = -f\left(t,x,u,a(t,x)\frac{\partial u}{\partial x}\right),$$

with the final condition $u(T,x) = \Phi(x)$.

Then if we put $Y_t = u(t,X_t)\,,\ Z_t = a(t,X_t)\,u'_x(t,X_t)$. By Itô's formula

$$\begin{aligned}\mathrm{d}Y_t &= \left[\frac{\partial u}{\partial t}(t,X_t) + b(t,X_t)\frac{\partial u}{\partial x}(t,X_t) + \frac{1}{2}a(t,x)^2\frac{\partial^2 u}{\partial x^2}(t,X_t)\right]\mathrm{d}t \\ &\qquad + a(t,X_t)\frac{\partial u}{\partial x}(t,X_t)\,\mathrm{d}W_t \\ &= -f(t,X_t,Y_t,Z_t)\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, \qquad Y_0 = u(0,X_0)\,.\end{aligned}$$

The final value $Y_T = u(T,X_T) = \Phi(X_T)$.

Statistical problems. We consider this problem in the situations, where the forward equation contains some unknown parameter ϑ:

$$\mathrm{d}X_t = b(\vartheta, t, X_t)\,\mathrm{d}t + a(\vartheta, t, X_t)\,\mathrm{d}W_t, \quad X_0 = x_0,\ 0 \le t \le T.$$

Then $u = u\,(t, x, \vartheta)$ and the proposed approximations $\hat{Y}_t, \hat{Z}_t$ of the couple Y_t, Z_t are given by the relations

$$\hat{Y}_t = u(t, X_t, \vartheta_t^*), \qquad\qquad \hat{Z}_t = u_x'(t, X_t, \vartheta_t^*)\,a(\vartheta_t^*, t, X_t).$$

Here ϑ_t^* is some *good* estimator-process of ϑ with the *small error*.

Our main problem is to find such estimator-processes $\hat{Y}_t, \hat{Z}_t, 0 < t \le T$ *that the errors* $\mathbf{E}_\vartheta \left(\hat{Y}_t - Y_t\right)^2$ and $\mathbf{E}_\vartheta \left(\hat{Z}_t - Z_t\right)^2$ *are minimal in the class of all possible estimators.*

Intermediary Problem: *How to find a good estimator-process* $\vartheta_t^*, 0 < t \leq T$? *Good* means :

- *It depends on observations* $X^t = (X_s, 0 \leq s \leq t)$ *and is stochastic process* $\vartheta^\star = \vartheta_t^\star, 0 < t \leq T$.
- *Easy to calculate for all* $t \in (0, T]$.
- *Asymptpotically efficient for all* $t \in (0, T]$ *(in some sense).*

The MLE-process $\left(\hat{\vartheta}_t, 0 < t \leq T\right)$ defined by the equation

$$V\left(\hat{\vartheta}_t, X^t\right) = \sup_{\vartheta \in \Theta} V\left(\vartheta, X^t\right), \quad 0 < t \leq T$$

can not be used as *good* because in non linear case to solve this equation for all $t \in (0, T]$ is a difficult problem.

We consider three forward diffusion processes in the situations where the consistent estimation is possible:

- Diffusion process with *small noise* $(\varepsilon \to 0)$

$$\mathrm{d}X_t = S\left(\vartheta, t, X_t\right)\mathrm{d}t + \varepsilon\sigma\left(t, X_t\right)\,\mathrm{d}W_t, \quad x_0,\ 0 \le t \le T.$$

- Discrete time observations $X^n = \left(X_{t_0}, X_{t_1}, \ldots X_{t_n}\right)$, $t_i = i\frac{T}{n}$ of the process $(n \to \infty)$

$$\mathrm{d}X_t = S\left(t, X_t\right)\mathrm{d}t + \sigma\left(\vartheta, t, X_t\right)\,\mathrm{d}W_t, \quad X_0,\ 0 \le t \le T.$$

- Ergodic diffusion process $(T \to \infty)$

$$\mathrm{d}X_t = S\left(\vartheta, X_t\right)\mathrm{d}t + \sigma\left(X_t\right)\,\mathrm{d}W_t, \quad X_0,\ 0 \le t \le T.$$

For each model we propose an estimator $\vartheta_t^\star = \vartheta_t^\star\left(X_s, 0 \le s \le t\right)$ such that $\hat{Y}_t = u(t, X_t, \vartheta_t^\star) \to Y_t$ and the error of approximation $\mathbf{E}_\vartheta\left(\hat{Y}_t - Y_t\right)^2$ is asymptotically optimal.

Kamatani and Uchida [6] recently considered the problem of parameter estimation by the discrete time observations of diffusion process and showed that multi-step Newton-Raphson procedure can provide asymptotically efficient estimation even if the preliminary estimators have bad rate of convergence. Our work is inspired by this result. We propose multi-step MLE-process with preliminary estimator constructed by a negligeable part of observations.

Example. *Ergodic diffusion.* Fix a learning interval $[0,\tau_T]$, where $\tau_T = T^{\delta} \to \infty, \delta < 1$ and obtain the preliminary estimator $\bar\vartheta_\tau$. Then we use this estimator to construct one-step $(\vartheta^\star_{t,T}, \tau_T \leq t \leq T)$ and two-step $(\vartheta^{\star,\star}_{t,T}, \tau_T \leq t \leq T)$ MLE-processes. Say, $(\tau = \tau_T)$

$$\vartheta^\star_{t,T} = \bar\vartheta_\tau + T^{-1}\mathbb{I}\left(\bar\vartheta_\tau\right)^{-1}\int_\tau^t \frac{\dot S\left(\bar\vartheta_\tau, X_s\right)}{\sigma\left(X_s\right)^2}\left[{\rm d}X_s - S\left(\bar\vartheta_\tau, X_s\right){\rm d}s\right].$$

This estimator-process is easy to calculate, it is uniformly on $\tau_T \leq t \leq T$ consistent, asymptotically normal and asymptotically efficient: if we put $t = sT$, then for any $s \in (0,1]$ as $T \to \infty$ we have

$$\sqrt{sT}\left(\vartheta^\star_{sT,T} - \vartheta\right) \Longrightarrow \mathcal{N}\left(0, \mathbb{I}\left(\vartheta\right)^{-1}\right).$$

The main contribution of this talk: the estimator-processess $(\hat Y_t, \hat Z_t, \tau_T < t \leq T)$ obtained with the help of $\vartheta^\star_{t,T}$ or $\vartheta^{\star,\star}_{t,T}$ are asymptotically efficient.

Example. *Time series.* (K. and Motrunich) Introduce the nonlinear AR process

$$X_j = X_{j-1} + 3\frac{\vartheta - X_{j-1}}{1 + \left(X_{j-1} - \vartheta\right)^2} + \varepsilon_j, \quad j = 1, \ldots, n,$$

where $\left(\varepsilon_j\right)_{j\geq 1}$ are i.i.d. $\mathcal{N}\left(0,1\right)$ and X_0 is given. The unknown parameter $\vartheta \in \Theta = (-1,1)$. We have to estimate ϑ and we need an estimator-process $\vartheta_n^* = \left(\vartheta_{j,n}^*, j = 1, \ldots, n\right)$, here $\vartheta_{j,n}^* = \vartheta_{j,n}^*(X_0^j)$.

We construct such process in two or three steps. First by $N = o\left(n\right)$ observations we obtain a preliminary consistent estimator $\bar\vartheta_N$ and then using this estimator we construct one or two step MLE-process as follows.

Case $N = n^{\delta}, \frac{1}{2} < \delta \leq 1$. Note that the unknown parameter is the shift parameter and that the invariant density function is symmetric with respect to ϑ. Hence we can put $N = \left[n^{3/4}\right]$ and

$$\bar{\vartheta}_N = \frac{1}{N}\sum_{j=1}^{N} X_j \longrightarrow \vartheta, \qquad n^{\frac{3}{8}}\left(\bar{\vartheta}_N - \vartheta\right) \Longrightarrow \mathcal{N}\left(0, \mathbb{B}\left(\vartheta\right)\right).$$

Of course, the limit variance of the EMM $\bar{\vartheta}_N$ is greater than that of the MLE, but this estimator is much more easier to calculate.

The score-function process is

$$\Delta_k(\vartheta, X^k) = \frac{1}{\sqrt{k}} \sum_{j=1}^{k} \dot{\ell}\left(\vartheta, X_{j-1}, X_j\right), \qquad N+1 \leq k \leq n.$$

where

$$\dot{\ell}\left(\vartheta, x, x'\right) = 3\left(x' - x - 3\frac{\vartheta - x}{1+\left(\vartheta - x\right)^2}\right) \frac{1-(\vartheta - x)^2}{\left(1+(\vartheta - x)^2\right)^2}.$$

Therefore we can calculate the one-step MLE-process as follows

$$\vartheta^\star_{k,n} = \bar\vartheta_N$$
$$+ \frac{3}{\mathbb{I}_k k} \sum_{j=1}^{k} \left(X_j - X_{j-1} - 3\frac{\bar\vartheta_N - X_{j-1}}{1+\left(\bar\vartheta_N - X_{j-1}\right)^2}\right) \frac{1-(\bar\vartheta_N - X_{j-1})^2}{\left(1+(\bar\vartheta_N - X_{j-1})^2\right)^2}$$

Here $\mathbb{I}_k$ is the empirical Fisher information.

If we put $k = [sn]\,, s \in (0,1]$ then

$$\sqrt{sn}\left(\vartheta^\star_{k,n} - \vartheta\right) \Longrightarrow \mathcal{N}\left(0, \mathbb{I}\left(\vartheta\right)^{-1}\right)$$

The one-step MLE-process admits the recurrent representation

$$\vartheta^\star_{k+1,n} = \frac{k\,\vartheta^\star_{k,n}}{k+1} + \frac{\bar\vartheta_N}{k+1} + \frac{1}{k+1}\mathbb{I}\left(\bar\vartheta_N\right)^{-1}\dot\ell\left(\bar\vartheta_N, X_k, X_{k+1}\right).$$

It allows us to calculate $\vartheta^\star_{k+1,n}$ using the values $\bar\vartheta_N, \vartheta^\star_{k,n}$ and observations X_k, X_{k+1} only.

If we decide that the leraning interval $\left[0,n^{\delta}\right]$ with $\delta\in(\frac{1}{2},1]$ is too long, then we can use shorter learning intervals, say, $\left[0,n^{\delta}\right]$ with $\delta\in(\frac{1}{4},\frac{1}{2}]$ or $\delta\in(\frac{1}{8},\frac{1}{4}]$ and then to construct the two-step or three-step MLE-processes respectively.

Case $N=\left[n^{\delta}\right], \frac{1}{4}<\delta\leq\frac{1}{2}$. The asymptotically efficient estimator we construct in three steps. By the first N observations as before we obtain the preliminary estimator $\bar\vartheta_N$ which is asymptotically normal with the rate $\sqrt{N}$, i.e.,

$$n^{\frac{\delta}{2}}\left(\bar\vartheta_N-\vartheta\right)\Longrightarrow\mathcal{N}\left(0,\mathbb{B}\left(\vartheta\right)\right).$$

This can be the same EMM.

The two-step MLE-process $\vartheta_n^{\star\star} = \left(\vartheta_{k,n}^{\star\star}, k = N+1, \ldots, n\right)$ we construct as follows. Introduce the second preliminary estimator-process

$$\bar\vartheta_{k,n} = \bar\vartheta_N + \frac{1}{\sqrt{k}} \mathbb{I}\left(\bar\vartheta_N\right)^{-1} \Delta_k(\bar\vartheta_N, X^k),$$

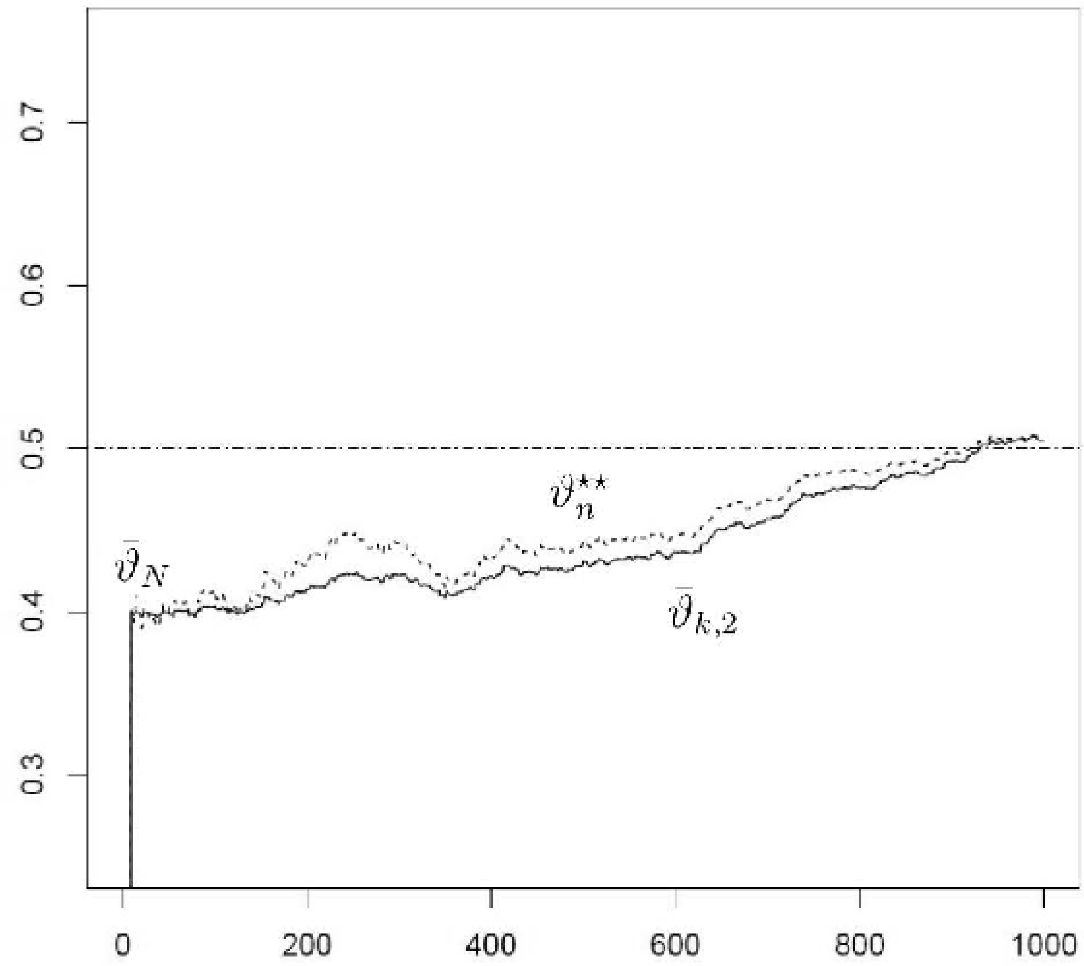
and two-step MLE-process

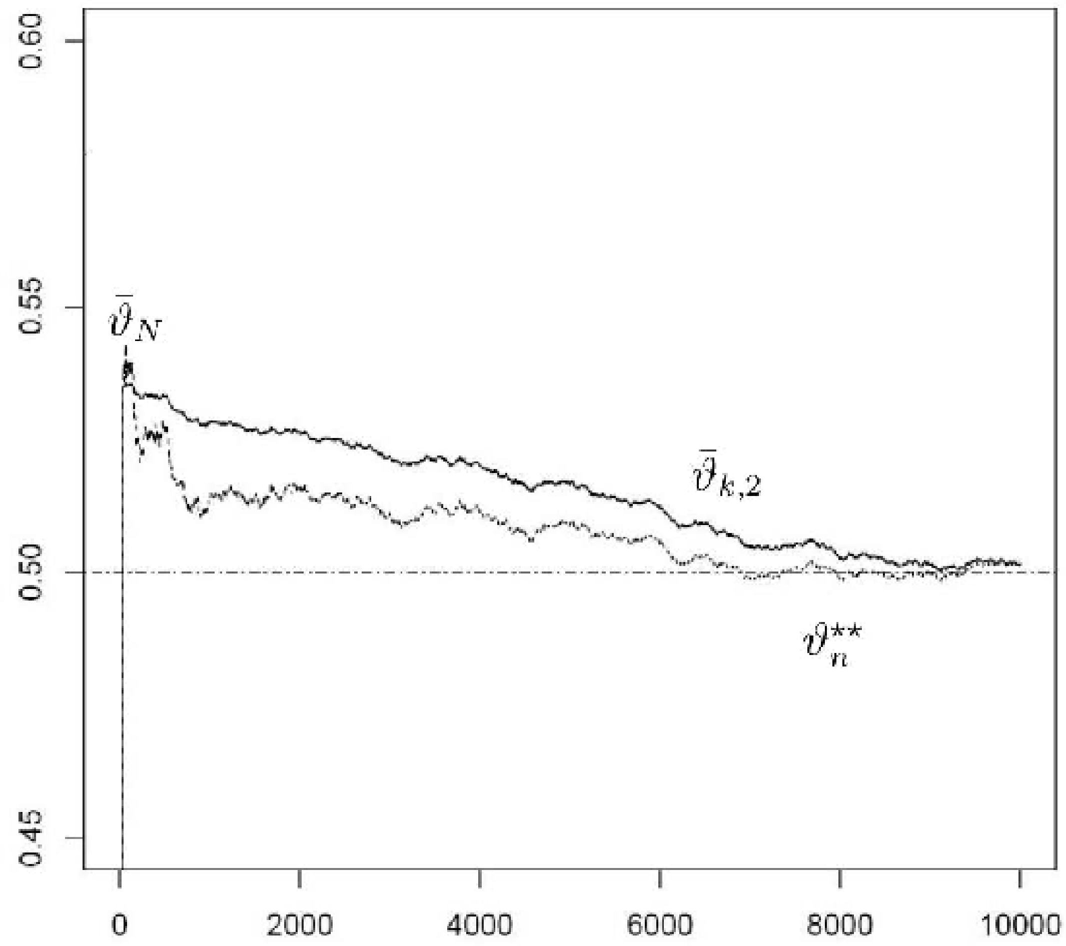
$$\vartheta_{k,n}^{\star\star} = \bar\vartheta_{k,n} + \frac{1}{\sqrt{k}} \mathbb{I}\left(\bar\vartheta_{k,n}\right)^{-1} \Delta_k(\bar\vartheta_{k,n}, X^k).$$

In the next theorem we realize this program.

Theorem 1 *Suppose that the conditions of regularity are fulfilled, then the estimator $\vartheta_{k,n}^{\star\star}$ is asymptotically normal*

$$\sqrt{sn}(\vartheta_{k,n}^{\star\star} - \vartheta) \Longrightarrow \mathcal{N}\left(0, \mathbb{I}\left(\vartheta\right)^{-1}\right).$$


0.7
0.6
0.5
0.4
0.3
0
200
400
600
800
1000
$\bar{\vartheta}_N$
$\vartheta_n^{\star\star}$
$\bar{\vartheta}_{k,2}$


$\bar{\vartheta}_N$
$\bar{\vartheta}_{k,2}$
$\vartheta_n^{\star\star}$
0.60
0.55
0.50
0.45
0
2000
4000
6000
8000
10000

Small noise asymptotics (joint work with L.Zhou)

The observed diffusion process (forward) is

$$\mathrm{d}X_t = S\left(\vartheta, t, X_t\right)\mathrm{d}t + \varepsilon\sigma\left(t, X_t\right)\,\mathrm{d}W_t, \quad X_0,\ 0 \leq t \leq T$$

where $\vartheta \in \Theta = (\alpha, \beta)$ is unknown parameter. We are given two functions $f\left(t, x, y, z\right)$, $\Phi\left(x\right)$ and we have to find a couple of stochastic processes $\left(\hat{X}_t, \hat{Z}_t, 0 \leq t \leq T\right)$ which approximate well the solution of the BSDE

$$\mathrm{d}Y_t = -f\left(t, X_t, Y_t, Z_t\right)\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, \qquad Y_0, \quad 0 \leq t \leq T$$

satisfying the condition $Y_T = \Phi\left(X_T\right)$. The functions $S\left(\cdot\right)$ and $\sigma\left(\cdot\right)$ are known and smooth. We have to minimize the errors

$$\mathbf{E}_\vartheta\left(\hat{Y}_t - Y_t\right)^2 \rightarrow \min, \qquad \mathbf{E}_\vartheta\left(\hat{Z}_t - Z_t\right)^2 \rightarrow \min$$

as $\varepsilon \rightarrow 0$.

Solution: Let us introduce a family of functions

$$\mathcal{U} = \left\{ \left(u(t,x,\vartheta), t \in [0,T], x \in \mathbb{R} \right), \vartheta \in \Theta \right\}$$

such that for all $\vartheta \in \Theta$ the function $u(t,x,\vartheta)$ satisfies the equation

$$\frac{\partial u}{\partial t} + S(\vartheta,t,x)\frac{\partial u}{\partial x} + \frac{\varepsilon^2 \sigma(t,x)^2}{2}\frac{\partial^2 u}{\partial x^2} = -f\left(t,x,u,\varepsilon\sigma(x)\frac{\partial u}{\partial x}\right)$$

and condition $u(T,x,\vartheta) = \Phi(x)$. If we put $Y_t = u(t, X_t, \vartheta)$, then by Itô's formula we obtain BSDE with $Z_t = \varepsilon\sigma(t, X_t)\, u'_x(t, X_t, \vartheta)$. As we do not know the value ϑ we propose first to estimate it using some estimator $\vartheta^\star_\varepsilon$ and then to put

$$\hat{Y}_t = u\left(t, X_t, \vartheta^\star_\varepsilon\right), \qquad \hat{Z}_t = \varepsilon\sigma\left(t, X_t\right) u'_x\left(t, X_t, \vartheta^\star_\varepsilon\right).$$

First we establish the low bounds on the risks of all estimator-processes and then show that the proposed estimator-processes attaint these bounds.

Theorem 2 *For all estimator-processes* $\bar{Y}_t, \bar{Z}_t, t \in [\tau, T]$ *we have the relations*

$$\underline{\lim}_{\nu \to 0} \underline{\lim}_{\varepsilon \to 0} \sup_{|\vartheta - \vartheta_0| \leq \nu} \varepsilon^{-2} \mathbf{E}_\vartheta \left| \bar{Y}_t - Y_t \right|^2 \geq \frac{\dot{u}^0 \left(t, x_t\left(\vartheta_0\right), \vartheta_0\right)^2}{\mathbb{I}_t\left(\vartheta_0, x^t\left(\vartheta_0\right)\right)},$$

$$\underline{\lim}_{\nu \to 0} \underline{\lim}_{\varepsilon \to 0} \sup_{|\vartheta - \vartheta_0| \leq \nu} \varepsilon^{-4} \mathbf{E}_\vartheta \left| \bar{Z}_t - Z_t \right|^2$$

$$\geq \frac{\left(\dot{u}^0\right)_x' \left(t, x_t\left(\vartheta_0\right), \vartheta_0\right)^2 \sigma\left(t, x_t\left(\vartheta_0\right)\right)^2}{\mathbb{I}_t\left(\vartheta_0, x^t\left(\vartheta_0\right)\right)}$$

Here $0 < \tau < T$ and $u^0\left(t, x_t\left(\vartheta_0\right), \vartheta_0\right)$ is solution of PDE for $\varepsilon = 0$.

We call an approximations $Y_t^\star, Z_t^\star, \tau \leq t \leq T$ asymptotically efficient if for all $\vartheta_0 \in \Theta$ and all $t \in [\tau, T]$ we have the equality

$$\lim_{\nu\to 0}\lim_{\varepsilon\to 0}\sup_{|\vartheta-\vartheta_0|\leq\nu}\varepsilon^{-2}\mathbf{E}_\vartheta\left|Y_t^\star - Y_t\right|^2 = \frac{\dot u^0\left(t, x_t\left(\vartheta_0\right), \vartheta_0\right)^2}{\mathbb{I}_t\left(\vartheta_0, x^t\left(\vartheta_0\right)\right)}$$

and

$$\lim_{\nu\to 0}\lim_{\varepsilon\to 0}\sup_{|\vartheta-\vartheta_0|\leq\nu}\varepsilon^{-4}\mathbf{E}_\vartheta\left|Z_t^\star - Z_t\right|^2$$

$$= \frac{\left(\dot u^0\right)_x'\left(t, x_t\left(\vartheta_0\right), \vartheta_0\right)^2\sigma\left(t, x_t\left(\vartheta_0\right)\right)^2}{\mathbb{I}_t\left(\vartheta_0, x^t\left(\vartheta_0\right)\right)}$$

Construction of the Estimator: Introduce a family of deterministic functions $\left\{\left(x_s\left(\vartheta\right),0\leq s\leq T\right),\vartheta\in\Theta\right\}$ solution of ODE

$$\frac{\mathrm{d}x_s}{\mathrm{d}s}=S\left(\vartheta,s,x_s\right),\qquad x_0,\quad 0\leq s\leq T.$$

It is known that X_s converges to $x_s\left(\vartheta\right)$ uniformly in $s\in[0,T]$.

Introduce the LR function

$$L\left(\vartheta,X^t\right)=\exp\left\{\int_0^t\frac{S\left(\vartheta,s,X_s\right)}{\varepsilon^2\,\sigma\left(s,X_s\right)^2}\,\mathrm{d}X_s-\int_0^t\frac{S\left(\vartheta,s,X_s\right)^2}{2\,\varepsilon^2\,\sigma\left(s,X_s\right)^2}\,\mathrm{d}s\right\}$$

and define the MLE-process $\hat\vartheta_{t,\varepsilon},0<t\leq T$ by the equation

$$L\left(\hat\vartheta_{t,\varepsilon},X^t\right)=\sup_{\vartheta\in\Theta}L\left(\vartheta,X^t\right).$$

It is known that $\varepsilon^{-1}\left(\hat\vartheta_{t,\varepsilon}-\vartheta_0\right)\Longrightarrow\mathcal{N}\left(0,\mathbb{I}_t\left(\vartheta,x^t\right)^{-1}\right)$, but to use it for $\bar Y_t=u\left(t,X_t,\hat\vartheta_{t,\varepsilon}\right)$ can be computantionally difficult problem.

Here the Fisher information $\mathbb{I}_t\left(\vartheta, x^t\left(\vartheta\right)\right)$ is

$$\mathbb{I}_t\left(\vartheta, x^t\left(\vartheta\right)\right) = \int_0^t \frac{\dot{S}\left(\vartheta, s, x_s\left(\vartheta\right)\right)^2}{\sigma\left(s, x_s\left(\vartheta\right)\right)^2}\, {\rm d}s.$$

As preliminary estimator we propose a MDE. Fix some (small) $\tau > 0$ and introduce the MDE $\bar{\vartheta}_{\tau,\varepsilon}$:

$$\left\|X - x\left(\bar{\vartheta}_{\tau,\varepsilon}\right)\right\|_\tau^2 = \inf_{\vartheta\in\Theta} \left\|X - x\left(\vartheta\right)\right\|_\tau^2 = \inf_{\vartheta\in\Theta} \int_0^\tau \left[X_t - x_t\left(\vartheta\right)\right]^2 \,{\rm d}t.$$

Suppose that the identifiability condition is fulfilled: for any $\nu > 0$

$$\inf_{\left|\vartheta-\vartheta_0\right|>\nu} \left\|x\left(\vartheta\right) - x\left(\vartheta_0\right)\right\|_\tau > 0.$$

This estimator is consistent and asymptotically normal

$$\varepsilon^{-1}\left(\bar{\vartheta}_{\tau,\varepsilon} - \vartheta_0\right) \Longrightarrow \mathcal{N}\left(0, \mathbb{D}_\tau\left(\vartheta_0\right)^2\right),$$

where $\mathbb{I}_\tau\left(\vartheta_0, x^\tau\left(\vartheta_0\right)\right) \geq D_\tau\left(\vartheta_0\right)^{-2} > 0$ (K. 1994).

Let us introduce the first *one-step MLE-process*

$$\vartheta^*_{t,\varepsilon} = \bar\vartheta_{\tau,\varepsilon} + \varepsilon \frac{\Delta_t\left(\bar\vartheta_{\tau,\varepsilon}, X_\tau^t\right) + \Delta_\tau\left(\bar\vartheta_{\tau,\varepsilon}, X^\tau\right)}{\mathbb{I}_t\left(\bar\vartheta_{\tau,\varepsilon}, x^t\left(\bar\vartheta_{\tau,\varepsilon}\right)\right)}, \qquad \tau \leq t \leq T,$$

where

$$\Delta_t\left(\vartheta, X_\tau^t\right) = \int_\tau^t \frac{\dot S\left(\vartheta, s, X_s\right)}{\varepsilon\sigma\left(s, X_s\right)^2}\left[\mathrm{d}X_s - S\left(\vartheta, s, X_s\right)\,\mathrm{d}s\right], \qquad t \in [\tau, T],$$

$$\Delta_\tau\left(\vartheta, X^\tau\right) = A\left(\vartheta, \tau, X_\tau\right) - \int_0^\tau A_s'\left(\vartheta, s, X_s\right)\,\mathrm{d}s$$

$$- \frac{\varepsilon^2}{2}\int_0^\tau B_x'\left(\vartheta, s, X_s\right)\sigma\left(s, X_s\right)^2\mathrm{d}s - \int_0^\tau \frac{\dot S\left(\vartheta, s, X_s\right) S\left(\vartheta, s, X_s\right)}{\sigma\left(s, X_s\right)^2}\mathrm{d}s,$$

$$B\left(\vartheta, s, x\right) = \frac{\dot S\left(\vartheta, s, x\right)}{\sigma\left(s, x\right)^2}, \qquad A\left(\vartheta, s, x\right) = \int_{x_0}^x B\left(\vartheta, s, z\right)\,\mathrm{d}z$$

It was shown that the approximations $\hat{Y}_t, \hat{Z}_t, \tau \leq t \leq T$, where

$$\hat{Y}_t = u\left(t, X_t, \vartheta^*_{t,\varepsilon}\right), \qquad \hat{Z}_t = \varepsilon\sigma\left(t, X_t\right) u'_x\left(t, X_t, \vartheta^*_{t,\varepsilon}\right)$$

are asymptotically efficient (K. and L. Zhou, 2014).

The goal of this work is to show that if we slightly modify the estimator $\vartheta^\star_{t,\varepsilon}$ then we obtain asymptotically efficient estimation of $Y_t, Z_t, \tau_\varepsilon \leq t \leq T$ even for some $\tau_\varepsilon \to 0$.

Introduce the additional condition: $\left|\dot{S}\left(\vartheta, x_0\right)\right| \geq \kappa > 0$ and put $\tau_\varepsilon = \varepsilon^\delta$, where $\delta \in (0,1)$. For the MLE $\hat{\vartheta}_{\tau_\varepsilon}$ we have

$$\frac{\sqrt{\tau_\varepsilon}}{\varepsilon}\left(\hat{\vartheta}_{\tau_\varepsilon} - \vartheta\right) \Longrightarrow \mathcal{N}\left(0, \frac{\sigma\left(x_0\right)^2}{\dot{S}\left(\vartheta, x_0\right)^2}\right)$$

Introduce another one-step MLE-process $\vartheta^\star_{t,\varepsilon}, \tau_\varepsilon \leq t \leq T$

$$\vartheta^\star_{t,\varepsilon} = \hat\vartheta_{\tau_\varepsilon} + \varepsilon \frac{\Delta_t\left(\bar\vartheta_{\tau_\varepsilon}, X_\tau^t\right)}{\mathbb{I}_t\left(\hat\vartheta_{\tau_\varepsilon}, x^t\left(\bar\vartheta_{\tau_\varepsilon}\right)\right)},$$

where

$$\Delta_t\left(\vartheta, X_\tau^t\right) = \int_\tau^t \frac{\dot S\left(\vartheta, s, X_s\right)}{\varepsilon\sigma\left(s, X_s\right)^2}\left[\mathrm{d}X_s - S\left(\vartheta, s, X_s\right)\,\mathrm{d}s\right], \quad t \in [\tau_\varepsilon, T].$$

We show that if $\delta < 1$ then

$$\varepsilon^{-1}\left(\vartheta^\star_{t,\varepsilon} - \vartheta\right) \Longrightarrow \mathcal{N}\left(0, \mathbb{I}_t\left(\vartheta, x^t\right)^{-1}\right)$$

Introduce the estimators

$$Y_t^\star = u\left(t, X_t, \vartheta^\star_{t,\varepsilon}\right), \qquad Z_t^\star = \varepsilon\sigma\left(t, X_t\right) u_x'\left(t, X_t, \vartheta^\star_{t,\varepsilon}\right)$$

Theorem 3 *Suppose the conditions of regularity hold, then the processes* $Y_t^\star, Z_t^\star, \tau_\varepsilon \leq t \leq T$ *have the representation*

$$Y_t^\star = Y_t + \varepsilon \dot u \left(t, X_t, \vartheta_0\right) \, \xi_t \left(\vartheta_0\right) + o\left(\varepsilon\right),$$

$$Z_t^\star = Z_t + \varepsilon^2 \sigma\left(t, X_t\right) \dot u'_x \left(t, X_t, \vartheta_0\right) \, \xi_t \left(\vartheta_0\right) + o\left(\varepsilon^2\right),$$

where

$$\xi_t\left(\vartheta_0\right) = \mathbb{I}_t\left(\vartheta, x^t\right)^{-1} \int_0^t \frac{\dot S\left(\vartheta, x_s\right)}{\sigma\left(x_s\right)} \mathrm{d}W_s$$

The random process $\eta_{t,\varepsilon} = \varepsilon^{-1}\left(Y_t^\star - Y_t\right), \tau \leq t \leq T$ *for any* $\tau \in (0, T]$ *converges in distribution to the process* $\xi_t\left(\vartheta_0\right), \tau \leq t \leq T$.

Let us show that the proposed approximations are asymptotically efficient.

Theorem 4 *The approximations*

$$Y_t^\star = u\left(t, X_t, \vartheta_{t,\varepsilon}^\star\right) \quad \text{and} \quad Z_t^\star = \varepsilon\sigma\left(t, X_t\right) u_x'\left(t, X_t, \vartheta_{t,\varepsilon}^\star\right)$$

for any $t \in (0, T]$ *are asymptotically efficient, i.e.,*

$$\lim_{\nu\to 0}\lim_{\varepsilon\to 0}\sup_{|\vartheta-\vartheta_0|\leq\nu} \varepsilon^{-2}\mathbf{E}_\vartheta \left|Y_t^\star - Y_t\right|^2 = \frac{\dot u^0\left(t, x_t\left(\vartheta_0\right), \vartheta_0\right)^2}{\mathbb{I}\left(\vartheta_0, x^t\left(\vartheta_0\right)\right)},$$

$$\lim_{\nu\to 0}\lim_{\varepsilon\to 0}\sup_{|\vartheta-\vartheta_0|\leq\nu} \varepsilon^{-4}\mathbf{E}_\vartheta \left|Z_t^\star - Z_t\right|^2 = \frac{\sigma\left(t, x_t\left(\vartheta_0\right)\right)^2 \left(\dot u^0\right)_x'\left(t, x_t, \vartheta_0\right)^2}{\mathbb{I}\left(\vartheta_0, x^t\left(\vartheta_0\right)\right)}$$

Unknown volatility (joint work with S. Gasparyan)

The forward equation is

$$\mathrm{d}X_t = S\left(t, X_t\right) \mathrm{d}t + \sigma\left(\vartheta, t, X_t\right)\, \mathrm{d}W_t, \quad X_0,\ 0 \leq t \leq T$$

where $\vartheta \in \Theta = (\alpha, \beta)$. We observe the solution of this equation in discrete times $t_i = i\frac{T}{n}$ and have to study the approximation $\hat{Y}_t = u(t, X_{t_k}, \hat{\vartheta}_{t_k}), k = 1, \ldots, n$, where k satisfies the conditions $t_k \leq t \leq t_{k+1}$ and the estimator $\hat{\vartheta}_{t_k}$ is construct by the observations $X^k = (X_0, X_{t_1}, \ldots, X_{t_k})$. Our goal is to realize the same program as above: we study the one-step pseudo-MLE, which can be relatively easy in calculation and has some properties of optimality.

On parameter estimation in diffusion coefficient. First of all remind that ϑ can be calculated without error if we have continuous time observations. To illustrate it we give one example.

Example. Suppose that $\sigma\left(\vartheta,t,x\right)=\sqrt{\vartheta}h\left(t,x\right),\vartheta\in\left(\alpha,\beta\right),\ \alpha>0$, and the observed process is

$$\mathrm{d}X_t=S\left(t,X_t\right)\mathrm{d}t+\sqrt{\vartheta}h\left(t,X_t\right)\mathrm{d}W_t,\quad X_0,\quad 0\leq t\leq T.$$

We suppose as well that $\int_0^t h\left(s,X_s\right)^2\mathrm{d}s>0$.

Let us write the Itô formula for X_t^2:

$$X_t^2=X_0^2+2\int_0^t X_s\,\mathrm{d}X_s+\vartheta\int_0^t h\left(s,X_s\right)^2\mathrm{d}s,\qquad 0\leq t\leq T.$$

Hence, for all $t\in(0,T]$ we have with probability 1 the equality

$$\hat\vartheta_t=\frac{X_t^2-X_0^2-2\int_0^t X_s\,\mathrm{d}X_s}{\int_0^t h\left(s,X_s\right)^2\mathrm{d}s}=\vartheta$$

The problem became more interesting if we consider the discrete time observations $X^n = \left(X_{t_1}, \ldots, X_{t_n}\right), t_j = j\frac{T}{n}$ and the problem of approximation in the *high frequency asymptotics* $(n \to \infty)$. Then in Example we obtain the estimator

$$\hat{\vartheta}_{t,k} = \frac{X_{t_k}^2 - X_0^2 - 2\sum_{j=1}^{k} X_{t_{k-1}}\left(X_{t_k} - X_{t_{k-1}}\right)}{\sum_{j=1}^{k} h\left(t_{j-1}, X_{t_{j-1}}\right)^2 \delta}, \quad \delta = \frac{T}{n}.$$

It can be easily shown that if $n \to \infty$ then we have $\hat{\vartheta}_{t,n} \to \vartheta$ and we can use it in the approximation of Y_t as follows $\hat{Y}_{t,n} = u(t, X_t, \hat{\vartheta}_{t,n})$. We can describe the distribution of error $\sqrt{n}\left(\hat{Y}_{t,n} - Y_t\right)$, but the estimator is not asymptotically optimal. We consider a different estimator.

Let us introduce the equation

$$X_{t_{j+1}} = X_{t_j} + S\left(t_j, X_{t_j}\right)\delta + \sigma\left(t_j, X_{t_j}, \vartheta\right)\left[W_{t_{j+1}} - W_{t_j}\right].$$

Note that conditional $(X_{t_0}, \ldots, X_{t_j})$ distribution

$$X_{t_{j+1}} - X_{t_j} - S\left(t_j, X_{t_j}\right)\delta \quad \sim \quad \mathcal{N}\left(0, \sigma\left(t_j, X_{t_j}, \vartheta\right)^2 \delta\right),$$

therefore we can itroduce the log pseudo-likelihood ratio

$$L\left(\vartheta, X^k\right) = -\frac{1}{2}\sum_{j=0}^{k-1} \ln\left[2\pi\sigma\left(t_j, X_{t_j}, \vartheta\right)^2 \delta\right] - \frac{1}{2}\sum_{j=0}^{k-1} \frac{\left(X_{t_{j+1}} - X_{t_j} - S\left(t_j, X_{t_j}\right)\delta\right)^2}{\sigma\left(t_j, X_{t_j}, \vartheta\right)^2 \delta}$$

The corresponding contrast function is

$$U_k\left(\vartheta, X^k\right) = \sum_{j=0}^{k-1} \ln a\left(t_j, X_{t_j}, \vartheta\right) + \sum_{j=0}^{k-1} \frac{\left(X_{t_{j+1}} - X_{t_j} - S\left(t_j, X_{t_j}\right)\delta\right)^2}{a\left(t_j, X_{t_j}, \vartheta\right)\delta}$$

where $a\left(t, x, \vartheta\right) = \sigma\left(t, x, \vartheta\right)^2$. The estimator $\hat{\vartheta}_{t,n}$ is define by

$$U_k\left(\hat{\vartheta}_{t,n}, X^k\right) = \inf_{\vartheta \in \Theta} U_k\left(\vartheta, X^k\right)$$

It is known that this estimator is consistent, asymptotically conditionally normal

$$\sqrt{n}\left(\hat{\vartheta}_{t,n} - \vartheta_0\right) \Longrightarrow \mathcal{N}\left(0, \mathbb{I}_t\left(\vartheta_0\right)^{-1}\right),$$

$$\mathbb{I}_t\left(\vartheta_0\right) = 2\int_0^t \frac{\dot{\sigma}\left(s, X_s, \vartheta_0\right)^2}{\sigma\left(s, X_s, \vartheta_0\right)^2}\, \mathrm{d}s$$

and asymptotically efficient (Dohnal(1987), Genon-Catalot, Jacod (1993)).

Note that the approximation $\hat{Y}_t = u(t, X_{t_k}, \hat{\vartheta}_{t,n})$ is computationally difficult to realize. That is why we propose as above the one-step pseudo-MLE. Let us fix some (small) $\tau \in (0, T)$. The PMLE estimator $\hat{\vartheta}_{\tau,n}$ constructed by $X_{t_{0,n}}, X_{t_{1,n}}, \ldots, X_{t_{N,n}}$, where N is chosen from the condition $t_{N,n} \leq \tau < t_{N+1,n}$, is consistent and asymptotically conditionally normal.

Introduce the normalized pseudo score-function and the empirical Fisher information

$$\Delta_{k,n}(\vartheta) = \sum_{j=0}^{k-1} \frac{\left[\left(X_{t_{j+1}} - X_{t_j} - S_j\ \delta\right)^2 - a_j(\theta)\,\delta\right] \dot{a}_j(\vartheta)}{2a_j(\vartheta)^2\sqrt{\delta}},$$

$$\mathbb{I}_{k,n}(\vartheta) = \frac{1}{2}\sum_{j=0}^{k-1} \frac{\dot{a}_j(\vartheta)^2}{a_j(\vartheta)^2}\ \delta = 2\sum_{j=0}^{k-1} \frac{\dot{\sigma}\left(t_j, X_{t_j}, \vartheta\right)^2}{\sigma\left(t_j, X_{t_j}, \vartheta\right)^2}\ \delta.$$

We have the stable convergence

$$\Delta_{k,n}\left(\vartheta_0\right) \Longrightarrow \sqrt{2}\int_0^t \frac{\dot\sigma\left(s,X_s,\vartheta_0\right)}{\sigma\left(s,X_s,\vartheta_0\right)}\,{\rm d}w_s$$

and the convergence in probability

$$\mathbb{I}_{k,n}\left(\vartheta_0\right) \to \mathbb{I}_t\left(\vartheta_0\right).$$

The approximation of the random function Y_t we will do with the help of the following one-step PMLE

$$\vartheta_{k,n}^\star = \hat\vartheta_N + \sqrt{\delta}\;\frac{\Delta_{k,n}(\hat\vartheta_N)}{{\rm I}_{k,n}(\hat\vartheta_N)}$$

and show that this estimator is asymptotically efficient and easy calculated for all $t\in\left[\tau,T\right]$ (or $N<k\leq n$).

We have the lower bound (Dohnal 87)

$$\underline{\lim}_{\gamma\to 0}\ \underline{\lim}_{n\to\infty}\ \sup_{|\vartheta-\vartheta_0|<\gamma} nT^{-1}\mathbf{E}_\vartheta\left(\bar\vartheta_{t,n}-\vartheta\right)^2 \geq \mathbf{E}_{\vartheta_0}\mathbb{I}_t\left(\vartheta_0\right)^{-1}.$$

The one-step PME is asymptotically efficient

$$\lim_{\gamma\to 0}\ \lim_{n\to\infty}\ \sup_{|\vartheta-\vartheta_0|<\gamma} nT^{-1}\mathbf{E}_\vartheta\left(\bar\vartheta_{t,n}-\vartheta\right)^2 = \mathbf{E}_{\vartheta_0}\mathbb{I}_t\left(\vartheta_0\right)^{-1}.$$

Introduce the estimators $Y^\star_{t_k,n}=u\left(t,X_{t_k},\vartheta^\star_{k,n}\right)$ and $Z^\star_{t_k,n}=u'_x\left(t,X_{t_k},\vartheta^\star_{k,n}\right)\sigma\left(t,X_{t_k},\vartheta^\star_{k,n}\right)$ of the random functions Y_t and Z_t respectively.

Theorem 5 *Suppose that the conditions of regularity hold, then the estimators* $\left(Y^\star_{t,n}, t\in[\tau,T]\right)$ *and* $\left(Z^\star_{t,n}, t\in[\tau,T]\right)$ *are consistent*

$$Y^\star_{t_k,n} \longrightarrow Y_t, \qquad Z^\star_{t_k,n} \longrightarrow Z_t,$$

and asymptotically conditionally normal (stable convergence)

$$\delta^{-1/2}\left(Y^\star_{t_k,n} - Y_{t_k}\right) \Longrightarrow \dot u\left(t, X_t, \vartheta_0\right)\ \xi_t\left(\vartheta_0\right),$$

$$\delta^{-1/2}\left(Z^\star_{t_k,n} - Z_{t_k}\right) \Longrightarrow \dot u'_x\left(t, X_t, \vartheta_0\right)\sigma\left(t, X_t, \vartheta_0\right)\ \ \xi_t\left(\vartheta_0\right)$$
$$+ u'_x\left(t, X_t, \vartheta_0\right)\dot\sigma\left(t, X_t, \vartheta_0\right)\ \ \xi_t\left(\vartheta_0\right),$$

The approximations $Y_t^\star$ and $Z_t^\star$ of the processes Y_t and Z_t are valid for the values $t \in [\tau, T]$. We take τ as a function of n, i.e., $\tau = \tau_n \to 0$. The rate of convergence of τ_n we take in such a way that the preliminary estimator $\hat{\vartheta}_{\tau_n}$ is still consistent and the one-step MLE $\vartheta_t^\star$ is asymptotically efficient.

Le us put $\tau_n = T/\ln n$. Then for $k = k_n \to \infty$ satisfying the condition $n^{-1} k_n \leq \tau_n < n^{-1} k_{n-1}$

Therefore, for the normalized the contrast-function we have the convergence

$$\tilde{U}_{k_n}\left(\vartheta, X^{k_n}\right) = \frac{U_{k_n}\left(\vartheta, X^{k_n}\right)}{\tau_n} \longrightarrow \ln a\left(0, x_0, \vartheta\right) + \frac{a\left(0, x_0, \vartheta_0\right)}{a\left(0, x_0, \vartheta\right)}.$$

Suppose that condition

$$\left|\frac{\dot\sigma\left(0,x_0,\vartheta\right)}{\sigma\left(0,x_0,\vartheta\right)}\right|\geq\kappa>0.$$

holds, then the estimator $\hat\vartheta_{\tau_n}$ defined with the help of this contrast function

$$\tilde U_{k_n}\left(\hat\vartheta_{\tau_n},X^{k_n}\right)=\inf_{\vartheta\in\Theta}\tilde U_{k_n}\left(\vartheta,X^{k_n}\right)$$

is consistent and asymptotically normal.

Introduce the one-step pseudo MLE

$$\vartheta_{k,n}^{\star}=\bar\vartheta_{\tau_n}+\sqrt{\delta}\,\frac{\Delta_{k,n}\left(\bar\vartheta_{\tau_n}\right)}{\mathrm{I}_{k,n}\left(\bar\vartheta_{\tau_n}\right)}.$$

This estimator is asymptotically efficient and easy calculated for all $N<k\leq n$.

Theorem 6 *Suppose that the conditions of regularity hold then*

$$\hat{Y}_{t,n} = u\left(t, X_{t_k}, \vartheta^\star_{k,n}\right) \longrightarrow Y_t,$$

$$\hat{Z}_{t,n} = u'_x\left(t, X_{t_k}, \vartheta^\star_{k,n}\right)\sigma\left(t, X_{t_k}, \vartheta^\star_{k,n}\right) \longrightarrow Z_t,$$

and the errors of estimation are

$$\delta^{-1/2}\left(\hat{Y}_{t_k,n} - Y_{t_k}\right) \Longrightarrow \dot{u}\left(t, X_t, \vartheta_0\right)\ \xi_t\left(\vartheta_0\right),$$

$$\delta^{-1/2}\left(\hat{Z}_{t_k,n} - Z_{t_k}\right) \Longrightarrow \left[\dot{u}'_x\left(t, X_t, \vartheta_0\right)\sigma\left(t, X_t, \vartheta_0\right) + u'_x\left(t, X_t, \vartheta_0\right)\dot{\sigma}\left(t, X_t, \vartheta_0\right)\right]\ \xi_t\left(\vartheta_0\right),$$

Observe that $Y_{t_k} - Y_t \sim O\left(\sqrt{\delta}\right)$. Below $\zeta \sim \mathcal{N}\left(0,1\right)$

$$\frac{\hat{Y}_{t_k,n} - Y_t}{\sqrt{\delta}} \Longrightarrow u'_x\left(t, X_t, \vartheta_0\right)\ \eta\,\sigma\left(t, X_t, \vartheta_0\right)\zeta + \dot{u}\left(t, X_t, \vartheta_0\right)\ \xi_t\left(\vartheta_0\right)$$

Example. The forward equation is

$$\mathrm{d}X_t = -X_t\mathrm{d}t + \sqrt{\theta + X_t^2}\,\mathrm{d}W_t, \quad X_0, \quad 0 \leq t \leq T.$$

Here $\vartheta \in \Theta = (\alpha, \beta)\,, \alpha > 0$ is unknown parameter. It is easy to see that in the case of continuous time observation the problem of parameter estimation is degenerated (singular), i.e., the unknown parameter ϑ can be estimated without error. Indeed, by Itô formula we can write

$$X_t^2 = X_0^2 + 2\int_0^t X_s\,\mathrm{d}X_s + \int_0^t \left[\vartheta + X_s^2\right]\mathrm{d}s.$$

Hence for all $t \in (0, T]$ we have the equality

$$\hat{\vartheta} = t^{-1}\left[X_t^2 - X_0^2 - 2\int_0^t X_s\,\mathrm{d}X_s - \int_0^t X_s^2\,\mathrm{d}s\right]$$

and $\hat{\vartheta} = \vartheta$

Our goal is to construct an asymptotically efficient estimator of the parameter ϑ. Note that the family of measures induced by the observations $X^k = (X_{t_0}, X_{t_1}, \ldots, X_{t_k})$ with t_k satisfying $t_k \leq t < t_{k+1}$ and fixed t are *locally asymptotically mixed normal* (LAMN) and for all estimators ϑ_k^* we have the lower bound on the risk

$$\underline{\lim}_{\nu\to 0}\, \underline{\lim}_{n\to\infty} \sup_{|\vartheta-\vartheta_0|<\nu} \mathbf{E}_\vartheta \ell\left(\sqrt{k}\left(\vartheta_k^* - \vartheta\right)\right) \geq \mathbf{E}_{\vartheta_0} \ell\left(\zeta_t\left(\vartheta_0\right)\right).$$

The first consistent estimator we obtain as follows

$$\bar\vartheta_N = \frac{n}{TN}\left[X_{t_N}^2 - X_0^2 - 2\sum_{j=1}^N X_{t_{j-1}}\left[X_{t_j} - X_{t_{j-1}}\right] - \sum_{j=1}^N X_{t_{j-1}}^2 \delta\right].$$

The pseudo log-likelihood ratio function is

$$L\left(\vartheta, X^N\right) = -\frac{1}{2}\sum_{j=1}^{N} \ln\left(2\pi\left(\vartheta + X_{t_{j-1}}^2\right)\right) - \sum_{j=1}^{N} \frac{\left[X_{t_j} - X_{t_{j-1}} + X_{t_{j-1}}\delta\right]^2}{2\left(\vartheta + X_{t_{j-1}}^2\right)\delta}.$$

Denote the pseudo Fisher information as

$$\mathbb{I}_{t_k,n}\left(\vartheta\right) = \frac{1}{2}\sum_{j=1}^{k} \frac{\delta}{\left(\vartheta + X_{t_{j-1}}^2\right)^2} \longrightarrow \mathbb{I}_t\left(\vartheta_0\right) = \frac{1}{2}\int_0^t \frac{\mathrm{d}s}{\left(\vartheta + X_s^2\right)^2}.$$

The one-step MLE-process $\vartheta^\star_{t_k,n}, \tau \leq t_k \leq T$ is

$$\vartheta^\star_{t_k,n} = \bar\vartheta_N + \sqrt{\delta}\sum_{j=1}^{k} \frac{\left[X_{t_j} - X_{t_{j-1}} + X_{t_{j-1}}\delta\right]^2 - \left(\bar\vartheta_N + X_{t_{j-1}}^2\right)\delta}{2\mathbb{I}_{t_k,n}\left(\bar\vartheta_N\right)\left(\bar\vartheta_N + X_{t_{j-1}}^2\right)^2\sqrt{\delta}}.$$

Example. Black-Scholes model. The forward equation is

$$\mathrm{d}X_t = \alpha X_t \mathrm{d}t + \vartheta X_t \,\mathrm{d}W_t, \quad X_0 = x_0, \quad 0 \le t \le T$$

and the function $f(x, y, z) = \beta y + \gamma x z$. The corresponding partial differential equation is

$$\frac{\partial u}{\partial t} + (\alpha + \vartheta\gamma)\, x \frac{\partial u}{\partial x} + \frac{\vartheta^2 x^2}{2} \frac{\partial^2 u}{\partial x^2} + \beta u = 0, \qquad u(T, x, \vartheta) = \Phi(x).$$

The solution of this equation is the function

$$u(t, x, \vartheta) = \frac{e^{\beta(T-t)}}{\sqrt{2\pi\vartheta^2 (T-t)}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2\vartheta^2 (T-t)}}\, \Phi\left(e^{x + \left(\alpha + \vartheta\gamma - \frac{\vartheta^2}{2}\right)(T-t) - z}\right) \mathrm{d}z.$$

The estimator of Y_t is

$$\hat{Y}_{t_k} = \int \frac{e^{-\frac{z^2}{2\hat{\vartheta}^2_{t_k,n}(T-t_k)}+\beta(T-t_k)}}{\sqrt{2\pi\hat{\vartheta}^2_{t_k,n}\,(T-t_k)}}\Phi\left(e^{X_{t_k}+(\alpha+\hat{\vartheta}_{t_k,n}\gamma-\frac{\hat{\vartheta}^2_{t_k,n}}{2})(T-t_k)-z}\right)\mathrm{d}z,$$

where $k = \left[\frac{t}{T}n\right]$ and

$$\hat{\vartheta}_{t_k,n} = \left(\frac{1}{t}\sum_{j=0}^{k-1}\frac{(X_{t_{j+1}} - X_{t_j} - \alpha X_{t_j}\delta)^2}{X^2_{t_j}}\right)^{\frac{1}{2}}.$$

Approximation of $\hat{Z}_t$.

Note that $u(t,x,\vartheta) = e^{\beta(T-t)}\mathbf{E}_{\vartheta,x}\Phi\left(e^{m_t-\xi}\right)$, where

$$\xi \sim \mathcal{N}\left(-x, d_t^2\right), \quad m_t = (\alpha + \vartheta\gamma - \frac{\vartheta^2}{2})\,(T-t)\,, \quad d_t^2 = \vartheta^2\,(T-t)$$

Hence

$$u_x'(t,x,\theta) = -e^{\beta(T-t)}\mathbf{E}_\theta\left[\frac{(x+\xi)}{d_t^2}\Phi(e^{m_t-\xi})\right]$$

and therefore

$$\hat{Z}_{t_k} = -\frac{\hat{\theta}_{t_k,n}\,X_{t_k}}{d_{t_k}^3\sqrt{2\pi}}\int_{-\infty}^{\infty}(y+X_{t_k})\Phi(e^{m_{t_k}-y})e^{-\frac{\left(X_{t_k}+y\right)^2}{2d_{t_k}^2}+\beta(T-t_k)}\mathrm{d}y$$

Ergodic diffusion (joint work with A. Abakirova)

The observed diffusion process (forward) is

$$\mathrm{d}X_t = S\left(\vartheta, X_t\right)\mathrm{d}t + \sigma\left(X_t\right)\,\mathrm{d}W_t, \quad X_0,\ 0 \leq t \leq T$$

where $\vartheta \in \Theta = (\alpha, \beta)$. The process $X_t, t \geq 0$ has ergodic properties. We are given two functions $f\left(x, y\right)$, $\Phi\left(x\right)$ and we have to find a couple of stochastic processes $\left(\hat{Y}_t, \hat{Z}_t, 0 \leq t \leq T\right)$ which approximate well the solution of the BSDE

$$\mathrm{d}Y_t = -f\left(X_t, Y_t, Z_t\right)\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, \qquad Y_0, \quad 0 \leq t \leq T$$

satisfying the condition $Y_T = \Phi\left(X_T\right)$. The functions $S\left(\cdot\right)$ and $\sigma\left(\cdot\right)$ are known and smooth. We have to minimize the errors

$$\mathbf{E}_\vartheta\left(\hat{Y}_t - Y_t\right)^2 \to \min, \qquad \mathbf{E}_\vartheta\left(\hat{Z}_t - Z_t\right)^2 \to \min.$$

as $T \to \infty$.

Solution: Introduce a family of functions
$\mathcal{U}=\{(u(t,x,\vartheta),t\in[0,T]\,,x\in\mathbb{R})\,,\vartheta\in\Theta\}$ such that for all $\vartheta\in\Theta$ the function $u(t,x,\vartheta)$ satisfies the equation

$$\frac{\partial u}{\partial t}+S(\vartheta,x)\frac{\partial u}{\partial x}+\frac{\sigma(x)^2}{2}\frac{\partial^2 u}{\partial x^2}=-f\left(x,u,\sigma\left(x\right)u_x'\right)$$

and condition $u(T,x,\vartheta)=\Phi\left(x\right)$. If we put $Y_t=u\left(t,X_t,\vartheta\right)$, then by Itô's formula we obtain BSDE with $Z_t=\sigma\left(X_t\right)u_x'\left(t,X_t,\vartheta\right)$.

Let us change the variables $t=sT, s\in[0,1]$, and put $v_\varepsilon\left(s,x,\vartheta\right)=u\left(sT,x,\vartheta\right)$, then

$$\varepsilon\frac{\partial v_\varepsilon}{\partial s}+S(\vartheta,x)\frac{\partial v_\varepsilon}{\partial x}+\frac{\sigma(x)^2}{2}\frac{\partial^2 v_\varepsilon}{\partial x^2}=-f\left(x,v_\varepsilon,\sigma\left(x\right)\left(v_\varepsilon\right)_x'\right),$$

where $v_\varepsilon\left(1,x,\vartheta\right)=\Phi\left(x\right)$ and $\varepsilon=T^{-1}$. The limit is $\varepsilon\to 0$.

We have a family of solutions $v_\varepsilon(s,y,\vartheta), 0\leq s\leq 1$. Fix some (small) $\delta>0$ and define the estimators

$$\hat Y_{sT}=v_\varepsilon\left(s,X_{sT},\vartheta^\star_{sT}\right),\qquad \hat Z_{sT}=\sigma\left(X_{sT}\right)\left(v_\varepsilon\right)'_x\left(s,X_{sT},\vartheta^\star_{sT}\right)$$

where $\vartheta^\star_{sT}, s\in[\delta,1]$ is one-step MLE, which is constructed as follows. Suppose that we have an estimator $\bar\vartheta_{\delta T}$ constructed by the observations $X^{\delta T}=(X_t, 0\leq t\leq \delta T)$, which is consistent and asymptotically normal

$$\sqrt{\delta T}\left(\bar\vartheta_{\delta T}-\vartheta\right)\Longrightarrow\mathcal{N}\left(0,D_\delta^2\right).$$

Then we calculate the one-step MLE

$$\vartheta^\star_{sT}=\vartheta^*_{\delta T}+\frac{\Delta_{sT}\left(\vartheta^*_{\delta T},X^{sT}_{\delta T}\right)+\Delta_\delta\left(\vartheta^*_{\delta T},X^{\delta T}\right)}{\sqrt{s\,T}\,\mathrm{I}\left(\vartheta^*_{\delta T}\right)},\qquad \delta\leq s\leq 1,$$

where

$$\Delta_{sT}\left(\vartheta, X_{\delta T}^{sT}\right) = \frac{1}{\sqrt{sT}} \int_{\delta T}^{sT} \frac{\dot S\left(\vartheta, X_t\right)}{\sigma\left(X_t\right)^2} \left[{\rm d}X_t - S\left(\vartheta, X_t\right)\,{\rm d}t\right], \qquad s \in [\delta, 1],$$

$$\Delta_{\delta}\left(\vartheta, X^{\delta T}\right) = \frac{A\left(\vartheta, X_\delta\right)}{\sqrt{sT}} - \frac{1}{2\sqrt{sT}} \int_0^\delta B_x'\left(\vartheta, X_t\right) \sigma\left(X_t\right)^2 {\rm d}t$$

$$- \int_0^\delta \frac{\dot S\left(\vartheta, X_t\right) S\left(\vartheta, X_t\right)}{\sqrt{sT}\sigma\left(X_t\right)^2} {\rm d}t,$$

$$B\left(\vartheta, x\right) = \frac{\dot S\left(\vartheta, x\right)}{\sigma\left(x\right)^2}, \qquad A\left(\vartheta, x\right) = \int_{x_0}^x B\left(\vartheta, z\right)\,{\rm d}z.$$

Note that under regularity conditions (K. 2004)

$$\sqrt{sT}\left(\vartheta_{sT}^\star - \vartheta\right) \Longrightarrow \mathcal{N}\left(0, \mathbb{I}\left(\vartheta\right)^{-1}\right)$$

$$\sqrt{sT}\left(\hat Y_{sT} - Y_{sT}\right) \sim \dot v_\varepsilon\left(s, X_{sT}, \vartheta\right) \sqrt{sT}\left(\vartheta_{sT}^\star - \vartheta\right),$$

$$\sqrt{sT}\left(\hat Z_{sT} - Z_{sT}\right) \sim \sigma\left(X_{sT}\right) \left(\dot v_\varepsilon\right)_x' \left(s, X_{sT}, \vartheta\right) \sqrt{sT}\left(\vartheta_{sT}^\star - \vartheta\right)$$

One-step MLE. Fix a learning interval $[0,\tau_T]$, where $\tau_T = T^\delta \to \infty, \delta \in (\frac{1}{2}, 1]$ and obtain the preliminary estimator $\bar\vartheta_\tau$. Then we construct one-step $(\vartheta^\star_{t,T}, \tau_T \leq t \leq T)$ and two-step $(\vartheta^{\star\star}_{t,T}, \tau_T \leq t \leq T)$ MLE-processes. Below $\tau = \tau_T$

$$\vartheta^\star_{t,T} = \bar\vartheta_\tau + T^{-1}\mathbb{I}\left(\bar\vartheta_\tau\right)^{-1}\int_\tau^t \frac{\dot S\left(\bar\vartheta_\tau, X_s\right)}{\sigma\left(X_s\right)^2}\left[{\rm d}X_s - S\left(\bar\vartheta_\tau, X_s\right){\rm d}s\right].$$

This estimator-process is easy to calculate, it is uniformly on $\tau_T \leq t \leq T$ consistent, asymptotically normal and asymptotically efficient: if we put $t = sT$, then for any $s \in (0,1]$ as $T \to \infty$ we have

$$\sqrt{sT}\left(\vartheta^\star_{sT,T} - \vartheta\right) \Longrightarrow \mathcal{N}\left(0, \mathbb{I}\left(\vartheta\right)^{-1}\right).$$

Two-step MLE. Let us take the *first* estimator $\tilde\vartheta_{\tau_\delta}$ constructed by the observations $X^{T^\delta} = \left(X_t, , 0 \leq t \leq T^\delta\right)$ with $\delta \in (\frac{1}{3}, \frac{1}{2}]$. We suppose that this estimator is consistent, asymptotically normal and the moments converge too:

$$\tilde v_{\tau_\delta} = T^{\frac{\delta}{2}} \left(\tilde\vartheta_{\tau_\delta} - \vartheta_0\right) \Longrightarrow \mathcal{N}\left(0, \mathbb{M}\left(\vartheta_0\right)\right), \qquad \sup_{\vartheta_0 \in \mathbb{K}} \mathbf{E}_{\vartheta_0} \left|\tilde v_{\tau_\delta}\right|^p \leq C,$$

for any $p > 0$. Here $\mathbb{M}\left(\vartheta_0\right)$ is some matrix and $C > 0$ does not depend on T. It can be the MLE, MDE, BE or the EMM (see [8]).

Introduce the *second* preliminary estimator-process

$$\bar\vartheta_\tau=\tilde\vartheta_{\tau_\delta}+\left(\tau T\right)^{-1/2}\mathbb{I}\left(\tilde\vartheta_{\tau_\delta}\right)^{-1}\Delta_{\tau T}\left(\tilde\vartheta_{\tau_\delta},X_{T^\delta}^{\tau T}\right),\qquad \tau\in\left[\tau_\delta,1\right]$$

where $\tau_\delta=T^{-1+\delta}$. Note that $T^\gamma\left(\bar\vartheta_\tau-\vartheta_0\right)\to 0$ for $\gamma\in\left(1-\delta,2\delta\right)$

$$\Delta_{\tau T}\left(\vartheta,X_{T^\delta}^{\tau T}\right)=\frac{1}{\sqrt{\tau T}}\int_{T^\delta}^{\tau T}\frac{\dot S\left(\vartheta,X_t\right)}{\sigma\left(X_t\right)^2}\left[{\rm d}X_t-S\left(\vartheta,X_t\right){\rm d}t\right].$$

Then the *two-step MLE-process* we define as follows

$$\vartheta_\tau^{\star\star}=\bar\vartheta_\tau+\frac{\mathbb{I}\left(\bar\vartheta_\tau\right)^{-1}}{\sqrt{\tau T}}\hat\Delta_{\tau T}\left(\tilde\vartheta_{\tau_\delta},\bar\vartheta_\tau,X_{T^\delta}^{\tau T}\right),\qquad \tau_\delta\leq\tau\leq 1,$$

where

$$\hat\Delta_{\tau T}\left(\vartheta_1,\vartheta_2,X_{T^\delta}^{\tau T}\right)=\frac{1}{\sqrt{\tau T}}\int_{T^\delta}^{\tau T}\frac{\dot S\left(\vartheta_1,X_t\right)}{\sigma\left(X_t\right)^2}\left[{\rm d}X_t-S\left(\vartheta_2,X_t\right){\rm d}t\right].$$

Note that $\hat\Delta_{\tau T}\left(\vartheta,\vartheta,X_{T^\delta}^{\tau T}\right)=\Delta_{\tau T}\left(\vartheta,X_{T^\delta}^{\tau T}\right)$.

Theorem 7 *Suppose that the conditions of regularity hold. Then the Two-step MLE-process* $\vartheta_\tau^{\star\star}, \tau_\delta \leq \tau \leq 1$ *is consistent, asymptotically normal*

$$\sqrt{T}\left(\vartheta_\tau^{\star\star} - \vartheta_0\right) \Longrightarrow \mathcal{N}\left(0, \tau^{-1}\mathbb{I}\left(\vartheta_0\right)^{-1}\right),$$

and asymptotically efficient. The random process

$$\eta_{\tau,T}\left(\vartheta_0\right) = \tau\sqrt{T}\mathbb{I}\left(\vartheta_0\right)^{-1/2}\left(\vartheta_\tau^{\star\star} - \vartheta_0\right), \qquad \tau_* \leq \tau \leq 1$$

for any $\tau_* \in (0,1)$ *converges in distribution to the Wiener process* $W\left(\tau\right), \tau_* \leq \tau \leq 1.$

References

[1] Abakirova, A. and Kutoyants Y.A. (2013) On approximation of the BSDE. Large samples approach. In progress.

[2] Dohnal, G. (1987) On estimating the diffusion coefficient. *J. Appl. Probab.*, 24, 1,105-114.

[3] Gasparyan, S. and Kutoyants, Y. (2015) On approximation of the BSDE with unknown volatility in forward equation. Armenian J. Mahemat., 7, 1, 59-79.

[4] Genon-Catalot, V. and Jacod, J. (1993) On the estimation of diffusion coefficient for multi-dimensional diffusion. *Ann. IHP*, Sec. B, 29, 1, 119-151.

[5] El Karoui N., Peng S. and Quenez M. (1997) Backward stochastic differential equations in finance, *Math. Fin.*, 7, 1-71.

[6] Kamatani, K. and Uchida, M. (2015) Hybrid multi-step

estimators for SDE based on sampled data. *Statist. Inference Stoch. Processes.* 18, 2, 177-204.

[7] Kutoyants Y.A. (1994) *Identification of Dynamical Systems with Small Noise,* Kluwer Academic Publisher, Dordrecht.

[8] Kutoyants Y.A. (2004) *Statiatical Inference for Ergodic Diffusion Processes.* Springer, London.

[9] Kutoyants Y.A. and Zhou, L.(2014) On approximation of the BSDE. (arXiv:1305.3728) *J. Stat. Plann. Infer.* , 150, 111-123.

[10] Pardoux E. and Peng S. (1990) Adapted solution of a BSDE. *System Control Letter*, 14, 55-61.

[11] Pardoux E. and Peng S. (1992) BSDE and quasilinear parabolic partial differential equations. *Stochastic Partial Differential Equations and their Applications* (Lect. Notes Control Inf. Sci. 176), 200-217, Springer, Berlin.