Hidden Markov Processes and Adaptive Filttering

Yury A. Kutoyants

Le Mans University, France

MASEMO workshop, Paris — July 1-4, 2025

1 Adaptive filtering

2 Different levels of noise

3 The same level of noise ($\psi_\varepsilon = \varepsilon$)

4 Low noise in observations ($\psi_\varepsilon = 1$)

5 Different Levels ($\psi_\varepsilon = \varepsilon^\delta, 1/3 < \delta < 1$)

6 Different Levels ($\delta = 1/3$)

7 Different Levels ($0 < \delta < 1/3$)

Kalman-Bucy filter We are given a couple of equations

$$
\begin{aligned}
\mathrm{d}X_t &= f\left(\vartheta, t\right) Y_t \mathrm{d}t + \sigma\left(t\right) \mathrm{d}W_t, \qquad X_0, \quad 0 \leq t \leq T,\\
\mathrm{d}Y_t &= a\left(\vartheta, t\right) Y_t \mathrm{d}t + b\left(\vartheta, t\right) \mathrm{d}V_t, \qquad Y_0 = y_0, \quad 0 \leq t \leq T,
\end{aligned}
$$

where $W_t, 0 \leq t \leq T$ and $V_t, 0 \leq t \leq T$ are independent Wiener processes and the observations are $X^T = \left(X_t, 0 \leq t \leq T\right)$. The process $Y^T = \left(Y_t, 0 \leq t \leq T\right)$ is *hidden*. The functions $f\left(\vartheta, t\right), \sigma\left(t\right), a\left(\vartheta, t\right), b\left(\vartheta, t\right), t \in [0, T]$ are supposed to be known, smooth and the parameter $\vartheta \in \Theta \subset \mathcal{R}^d$ is unknown. The true value is ϑ_0 and the initial value y_0 is deterministic and known.

The conditional expectation $m(\vartheta,t)=\mathbf{E}_\vartheta\left(Y_t|X_s,0\leq s\leq t\right)$ and the error $\gamma(\vartheta,t)=\mathbf{E}_\vartheta\left(Y_t-m(\vartheta,t)\right)^2$ are solutions of Kalman-Bucy (K-B) filtration equations (Kalman and Bucy, 1961)

$$\begin{aligned}\mathrm{d}m(\vartheta,t)&=\left[a(\vartheta,t)-\frac{\gamma(\vartheta,t)\,f(\vartheta,t)^2}{\sigma(t)^2}\right]m(\vartheta,t)\,\mathrm{d}t\\&\qquad+\frac{\gamma(\vartheta,t)\,f(\vartheta,t)}{\sigma(t)^2}\mathrm{d}X_t,\\\frac{\partial\gamma(\vartheta,t)}{\partial t}&=2a(\vartheta,t)\,\gamma(\vartheta,t)-\frac{\gamma(\vartheta,t)^2\,f(\vartheta,t)^2}{\sigma(t)^2}+b(\vartheta,t)^2,\end{aligned}$$

with initial values $m_0=\mathbf{E}_\vartheta\left(Y_0|X_0\right)$ and $\gamma_0=\mathbf{E}_\vartheta\left(Y_0-m_0\right)^2$. Recall that $m(\vartheta,t)$ is an optimal estimator of Y_t.

Pb.: *To obtain a good recursive approximation $m_t^\star, 0<t\leq T$ of the process $m(\vartheta_0,t)\,,0<t\leq T$.*

How to chose an estimator? The likelihood ratio function is:

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

The MLE $\hat{\vartheta}_t$ and BE $\tilde{\vartheta}_t$ are defined by the relations

$$L(\hat{\vartheta}_t, X^t) = \sup_{\vartheta\in\Theta} L\left(\vartheta, X^t\right), \qquad \tilde{\vartheta}_t = \frac{\int_\Theta \vartheta p\left(\vartheta\right) L\left(\vartheta, X^t\right)\mathrm{d}\vartheta}{\int_\Theta p\left(\vartheta\right) L\left(\vartheta, X^t\right)\mathrm{d}\vartheta}.$$

To construct the MLE and BE of the parameter ϑ we have to calculate the functions $\left\{m\left(\vartheta,s\right),\ 0\leq s\leq t\right\},\ \vartheta\in\Theta$ and $\left\{\gamma\left(\vartheta,s\right),\ 0\leq s\leq t\right\},\ \vartheta\in\Theta$ defined by the K-B equations.

Remind that but we can not put the MLE $\hat{\vartheta}_T$ or BE $\tilde{\vartheta}_T$ in $m\left(\vartheta,t\right)$ because the stochastic integral containing $\gamma(\hat{\vartheta}_T, t)$ does not defined.

Remark that the direct numerical calculations of the MLE $\hat\vartheta_t, 0 < t \leq T$ or BE $\tilde\vartheta_t, 0 < t \leq T$ are almost impossible.

To approximate $m(\vartheta_0, t)$ we propose the following program:

1. *Calculate a preliminary estimator* $\bar\vartheta_\tau$ *on relatively small interval of observations* $[0, \tau]$ *(*$\tau/T \to 0$*).*
2. *Using* $\bar\vartheta_\tau$ *realize One-step MLE-process* $\vartheta^\star_t, \tau < t \leq T$
3. *As approximation of* $m(\vartheta_0, t)$ *we propose* $m^\star_t$ *obtained with the help of K-B equations, where* ϑ *is replaced by* $\vartheta^\star_t, \tau < t \leq T$
4. *Estimate the error* $m^\star_t - m(\vartheta_0, t)$ *(asymptotic efficiency?).*

We apply this construction and propose adaptive filters to different models of observations in K. "*Hidden Markov Processes and Adaptive Filtering*", Springer series in Statistics, 2025.

1. Small noise in both equations: $\sigma(t) \to \varepsilon\sigma(t)$, $b(\vartheta, t) \to \psi_\varepsilon b(\vartheta, t)$, $\varepsilon \to 0, \psi_\varepsilon \to 0$
2. Small noise in observations only: $\sigma(t) \to \varepsilon\sigma(t)$
3. Conditionally Gaussian processes: $f(\vartheta, t, X_t)\, Y_t, a(\vartheta, t, X_t)\, Y_t$.
4. EKF: $f(\vartheta, t, Y_t)$ etc. $\varepsilon \to 0$
5. Hidden O-U process: $f, \sigma, -a, b,\ T \to \infty$
6. Hidden telegraph signal: $Y_t = a$ or $Y_t = b,\ T \to \infty$
7. Hidden AR process: $X_t = fY_t + \sigma w_t,\ Y_t = aY_{t-1} + bv_t, T \to \infty$
8. Several applications (localization of the fixed and moving source on the plane, identification of security price process, change-point problems, BSDE approximation)

1. Adaptive filtering
2. Different levels of noise
3. The same level of noise ($\psi_\varepsilon = \varepsilon$)
4. Low noise in observations ($\psi_\varepsilon = 1$)
5. Different Levels ($\psi_\varepsilon = \varepsilon^\delta, 1/3 < \delta < 1$)
6. Different Levels ($\delta = 1/3$)
7. Different Levels ($0 < \delta < 1/3$)

Basic Model

Suppose that

$$
\begin{aligned}
\mathrm{d}X_t &= f(\vartheta,t)\,Y_t\,\mathrm{d}t + \varepsilon\,\sigma(t)\,\mathrm{d}W_t, \qquad X_0 = 0, \quad 0 \le t \le T,\\
\mathrm{d}Y_t &= a(\vartheta,t)\,Y_t\,\mathrm{d}t + \psi_\varepsilon\, b(\vartheta,t)\,\mathrm{d}V_t, \qquad Y_0 = y_0.
\end{aligned}
$$

Observations are $X^T = (X_t, 0 \le t \le T)$ and the asymptotic $\varepsilon \to 0$.

The properties of MLEs are defined by the asymptotic of the integral $(M(\vartheta,t) = f(\vartheta,t)\,m(\vartheta,t))$

$$
\int_0^T \frac{\left[M(\vartheta_0 + \varphi_\varepsilon u, t) - M(\vartheta_0,t)\right]^2}{\varepsilon^2\,\sigma(t)^2}\mathrm{d}t \longrightarrow u^\top \mathbf{I}(\vartheta_0)\,u.
$$

Set $\psi_\varepsilon = \varepsilon^\delta, \delta \in [0,1]$ and $\varphi_\varepsilon = \varepsilon^{h(\delta)}$. Then in regular problems

$$
\varphi_\varepsilon^{-1}\left(\hat\vartheta_\varepsilon - \vartheta_0\right) \Longrightarrow \mathcal{N}\left(0, \mathbf{I}(\vartheta_0)^{-1}\right).
$$

Denote $a^*(\vartheta,t)=a(\vartheta,t)+\ell(\vartheta,t)$, $\ell(\vartheta,t)=\left[\ln f(\vartheta,t)\right]'_t$. The key representation is

$$M(\vartheta,t)-M(\vartheta_0,t)=\frac{\varepsilon}{\psi_\varepsilon}\frac{a^*(\vartheta,t)-a^*(\vartheta_0,t)}{B(\vartheta,t)}f(\vartheta_0,t)\,y_t(\vartheta_0)\,(1+o(1)$$

$$+\sqrt{\varepsilon\psi_\varepsilon}\,\frac{\left[f(\vartheta,t)\,b(\vartheta,t)-f(\vartheta_0,t)\,b(\vartheta_0,t)\right]}{\sqrt{2B(\vartheta,t)}}\;\;\xi_{t,\varepsilon}\,(1+o(1))\,,$$

where $B(\vartheta,t)=b(\vartheta,t)\,\sigma(t)\,f(\vartheta,t)^{-1}$ and

$$\xi_{t,\varepsilon}\Longrightarrow\xi_t\;\sim\;\mathcal{N}(0,1)\,.$$

The random variables $\xi_t, 0<t\leq T$ are independent.

Then we obtain 5 different problems:

1. $\psi_\varepsilon = \varepsilon$, $(\delta = 1)$, $\quad \varphi_\varepsilon = \varepsilon$,
2. $\psi_\varepsilon = 1$, $(\delta = 0)$, $\quad \varphi_\varepsilon = \sqrt{\varepsilon}$,
3. $\psi_\varepsilon = \varepsilon^\delta$, where $1/3 < \delta < 1$, $\quad \varphi_\varepsilon = \varepsilon^\delta$,
4. $\psi_\varepsilon = \varepsilon^{1/3}$, $\quad \varphi_\varepsilon = \varepsilon^{\frac{1}{3}}$,
5. $\psi_\varepsilon = \varepsilon^\delta$, where $0 < \delta < 1/3$, $\quad \varphi_\varepsilon = \varepsilon^{\frac{1-\delta}{2}}$.

$$\varphi_\varepsilon = \varepsilon^{h(\delta)}.$$

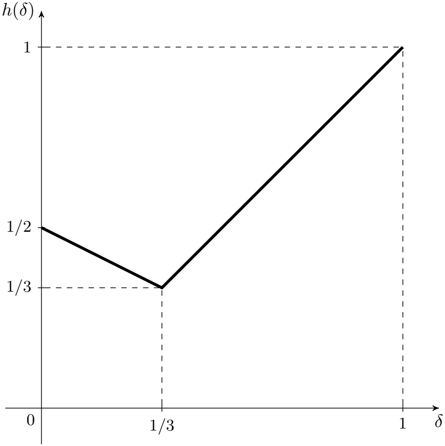

Figure: Function $h(\delta)$

For $\delta \in \left(0, \frac{1}{3}\right)$ the higher noise $\longrightarrow$ the smaller error of estimation!

Remind our program

- Preliminary consistent estimator $\bar\vartheta_{\tau,\varepsilon}$ by observations $X_t, 0 \leq t \leq \tau$.
- Properties of the MLE $\hat\vartheta_\varepsilon$ and BE $\tilde\vartheta_\varepsilon$.
- One-step MLE-process $\vartheta^\star_{t,\varepsilon}, \tau < t \leq T$
- Adaptive filter $m^\star_{t,\varepsilon}, \tau < t \leq T$
- Asymptotically efficient adaptive filter.

1 Adaptive filtering

2 Different levels of noise

3 The same level of noise ($\psi_\varepsilon = \varepsilon$)

4 Low noise in observations ($\psi_\varepsilon = 1$)

5 Different Levels ($\psi_\varepsilon = \varepsilon^\delta, 1/3 < \delta < 1$)

6 Different Levels ($\delta = 1/3$)

7 Different Levels ($0 < \delta < 1/3$)

Consider the linear two-dimensional partially observed system

$$\begin{aligned} \mathrm{d}X_t &= f(\vartheta, t)\, Y_t \mathrm{d}t + \varepsilon\sigma(t)\, \mathrm{d}W_t, \qquad X_0 = 0, \quad 0 \le t \le T, \\ \mathrm{d}Y_t &= a(\vartheta, t)\, Y_t \mathrm{d}t + \varepsilon b(t)\, \mathrm{d}V_t, \quad Y_0 = y_0 \ne 0. \end{aligned}$$

Asymptotic $\varepsilon \to 0$. The limit ($\varepsilon = 0$, ϑ_0 is the true value) system is

$$\begin{aligned} x_t(\vartheta_0) &= \int_0^t f(\vartheta_0, s)\, y_s(\vartheta_0)\, \mathrm{d}s, \qquad 0 \le t \le T \\ y_t(\vartheta_0) &= y_0 \exp\left(\int_0^t a(\vartheta_0, s)\, \mathrm{d}s\right). \end{aligned}$$

MDE. Set $[0, \tau_\varepsilon]$ (learning interval, $\tau_\varepsilon = \varepsilon^{2/3} \to 0$) and by observations $X^{\tau_\varepsilon} = (X_t, 0 \leq t \leq \tau_\varepsilon)$ we construct the MDE

$$\check{\vartheta}_{\tau_\varepsilon} = \arg \inf_{\vartheta \in \Theta} \int_0^{\tau_\varepsilon} \left[X_t - x_t\left(\vartheta\right)\right]^2 \mathrm{d}t$$

Regularity conditions: the functions $f\left(\vartheta, t\right), a\left(\vartheta, t\right)$ have two continuous derivatives on ϑ, $\left|\dot{f}\left(\vartheta, 0\right)\right| > 0$. It is shown that

Proposition

The MDE is consistent, as. normal

$$\frac{\tau_\varepsilon^{1/2}}{\varepsilon}\left(\check{\vartheta}_{\tau_\varepsilon} - \vartheta_0\right) \Longrightarrow \mathcal{N}\left(0, D\left(\vartheta_0\right)\right), \qquad D\left(\vartheta_0\right) = \frac{6}{5}\frac{\sigma\left(0\right)^2}{y_0 \dot{f}\left(\vartheta_0, 0\right)^2},$$

and the moments converge.

Here $\tau_\varepsilon^{1/2}\varepsilon^{-1} = \varepsilon^{-2/3}$.

The system can be simplified if we denote $\gamma(\vartheta, t) = \varepsilon^2 \gamma_*(\vartheta, t)$.

$$\mathrm{d}m(\vartheta, t) = \left[a(\vartheta, t) - \frac{\gamma_*(\vartheta, t) f(\vartheta, t)^2}{\sigma(t)^2}\right] m(\vartheta, t)\,\mathrm{d}t + \frac{\gamma_*(\vartheta, t) f(\vartheta, t)}{\sigma(t)^2}\mathrm{d}X_t,$$

$$\frac{\partial \gamma_*(\vartheta, t)}{\partial t} = 2a(\vartheta, t)\gamma_*(\vartheta, t) - \frac{\gamma_*(\vartheta, t)^2 f(\vartheta, t)^2}{\sigma(t)^2} + b(\vartheta, t)^2 .$$

Denote $y_t(\vartheta, \vartheta_0) = m(\vartheta, t)\,|_{\varepsilon=0}$. It is solution of the equation

$$\frac{\mathrm{d}y_t(\vartheta, \vartheta_0)}{\mathrm{d}t} = \left[a(\vartheta, t) - \frac{\gamma_*(\vartheta, t) f(\vartheta, t)^2}{\sigma(t)^2}\right] y_t(\vartheta, \vartheta_0, t) + \frac{\gamma_*(\vartheta, t) f(\vartheta, t) f(\vartheta_0, t)}{\sigma(t)^2} y_t(\vartheta_0) .$$

Introduce the notation: $S_*(\vartheta, \vartheta_0, t) = f(\vartheta, t)\, y_t(\vartheta, \vartheta_0)$,

$$\dot{S}_*(\vartheta, \vartheta_0, t) = \dot{f}(\vartheta, t)\, y_t(\vartheta, \vartheta_0) + f(\vartheta, t)\, \dot{y}_t(\vartheta, \vartheta_0),$$

$$\mathbf{I}(\vartheta_0) = \int_0^T \frac{\dot{S}_*(\vartheta_0, \vartheta_0, t)^2}{\sigma(t)^2}\mathrm{d}t.$$

Here $\mathbf{I}(\vartheta_0), \vartheta_0 \in \Theta$ is the Fisher information.

Theorem

The MLE and BE are uniformly on compacts $\mathbb{K} \subset \Theta$ consistent, asymptotically normal

$$\frac{\hat{\vartheta}_\varepsilon - \vartheta_0}{\varepsilon} \Longrightarrow \zeta(\vartheta_0) \sim \mathcal{N}\left(0, \mathbf{I}(\vartheta_0)^{-1}\right), \qquad \frac{\tilde{\vartheta}_\varepsilon - \vartheta_0}{\varepsilon} \Longrightarrow \zeta(\vartheta_0)$$

and asymptotically efficient.

Le Cam's **One-step MLE** $\vartheta_\varepsilon^\star$ is

$$\vartheta_\varepsilon^\star = \check\vartheta_{\tau_\varepsilon} + \mathbf{I}_{\tau_\varepsilon}^{T}(\check\vartheta_{\tau_\varepsilon})^{-1}\int_{\tau_\varepsilon}^{T}\frac{\dot M(\check\vartheta_{\tau_\varepsilon},s)}{\sigma\left(s\right)^2}\left[\mathrm{d}X_s - M(\check\vartheta_{\tau_\varepsilon},s)\mathrm{d}s\right],$$

and **One-step MLE-process** $\vartheta_{t,\varepsilon}^\star, \tau_\varepsilon < t \le T$

$$\vartheta_{t,\varepsilon}^\star = \check\vartheta_{\tau_\varepsilon} + \mathbf{I}_{\tau_\varepsilon}^{t}(\check\vartheta_{\tau_\varepsilon})^{-1}\int_{\tau_\varepsilon}^{t}\frac{\dot S_*(\check\vartheta_{\tau_\varepsilon},\check\vartheta_{\tau_\varepsilon},s)}{\sigma\left(s\right)^2}\left[\mathrm{d}X_s - M(\check\vartheta_{\tau_\varepsilon},s)\mathrm{d}s\right].$$

Here $M\left(\vartheta,s\right) = f\left(\vartheta,s\right)m\left(\vartheta,s\right)$ and the K-B equation for $M(\check\vartheta_{\tau_\varepsilon},s), \tau_\varepsilon \le s \le T$ is $\left(a^*\left(\vartheta,t\right) = a\left(\vartheta,t\right) + \ell\left(\vartheta,t\right)\right)$

$$\begin{aligned}\mathrm{d}M(\check\vartheta_{\tau_\varepsilon},s) = &\left[a^*(\check\vartheta_{\tau_\varepsilon},t) - \gamma_*(\check\vartheta_{\tau_\varepsilon},t)f(\check\vartheta_{\tau_\varepsilon},t)^2\sigma(t)^{-2}\right]M(\check\vartheta_{\tau_\varepsilon},t)\mathrm{d}t\\ &+ \gamma_*(\check\vartheta_{\tau_\varepsilon},t)f(\check\vartheta_{\tau_\varepsilon},t)^2\sigma\left(t\right)^{-2}\mathrm{d}X_t\end{aligned}$$

Fisher information function

$$\mathbf{I}_\tau^t(\vartheta_0) = \int_\tau^t \frac{\dot{S}_*(\vartheta_0, \vartheta_0, s)^2}{\sigma(s)^2} \mathrm{d}s > 0, \quad \tau < t \le T, \quad \vartheta \in \Theta.$$

Proposition

This estimator-process for all $t \in (\tau_\tau, T]$ is consistent, asymptotically normal

$$\frac{\vartheta_{t,\varepsilon}^\star - \vartheta_0}{\varepsilon} \Longrightarrow \zeta_t^\star(\vartheta_0) \sim \mathcal{N}\left(0, \mathbf{I}_0^t(\vartheta_0)^{-1}\right).$$

the moments converge and it is asymptotically efficient.

The estimator $\check\vartheta_{\tau_\varepsilon}$ consistent and as. norm.

$$\frac{\vartheta^\star_{t,\varepsilon}-\vartheta_0}{\varepsilon}=\frac{\check\vartheta_{\tau_\varepsilon}-\vartheta_0}{\varepsilon}+\mathbf{I}^t_{\tau_\varepsilon}(\check\vartheta_{\tau_\varepsilon})^{-1}\int_{\tau_\varepsilon}^{t}\frac{\dot S_*(\check\vartheta_{\tau_\varepsilon},\check\vartheta_{\tau_\varepsilon},s)}{\sigma\left(s\right)}\mathrm{d}\bar W_s$$
$$+\mathbf{I}^t_{\tau_\varepsilon}(\check\vartheta_{\tau_\varepsilon})^{-1}\int_{\tau_\varepsilon}^{t}\frac{\dot S_*(\check\vartheta_{\tau_\varepsilon},\check\vartheta_{\tau_\varepsilon},s)}{\varepsilon\sigma\left(s\right)^2}\left[M(\vartheta_0,s)-M(\check\vartheta_{\tau_\varepsilon},s)\right]\mathrm{d}s.$$

We have

$$\mathbf{I}^t_{\tau_\varepsilon}(\check\vartheta_{\tau_\varepsilon})^{-1}\int_{\tau_\varepsilon}^{t}\frac{\dot S_*(\check\vartheta_{\tau_\varepsilon},\check\vartheta_{\tau_\varepsilon},s)}{\sigma\left(s\right)}\mathrm{d}\bar W_s\Longrightarrow\zeta^\star_t\left(\vartheta_0\right)\sim\mathcal{N}\left(0,\mathbf{I}^t_0\left(\vartheta\right)^{-1}\right)$$

Taylor expansion of $M(\vartheta_0, s)$ at the point $\check{\vartheta}_{\tau_\varepsilon}$

$$\frac{(\vartheta_0 - \check{\vartheta}_{\tau_\varepsilon})}{\varepsilon} \frac{1}{\mathbf{I}_{\tau_\varepsilon}^t(\check{\vartheta}_{\tau_\varepsilon})} \int_{\tau_\varepsilon}^t \frac{\dot{S}_*(\check{\vartheta}_{\tau_\varepsilon}, \check{\vartheta}_{\tau_\varepsilon}, s)^2}{\sigma(s)^2} \mathrm{d}s + \frac{(\check{\vartheta}_{\tau_\varepsilon} - \vartheta_0)^2}{\varepsilon} O(1).$$

Set $\tau_\varepsilon = \varepsilon^{2/3}$

$$\begin{aligned} \frac{\left(\check{\vartheta}_{\tau_\varepsilon} - \vartheta_0\right)^2}{\varepsilon} &= \left[\frac{\tau_\varepsilon^{1/2}}{\varepsilon}\left(\check{\vartheta}_{\tau_\varepsilon} - \vartheta_0\right)\right]^2 \frac{\varepsilon^2}{\tau_\varepsilon\,\varepsilon} \\ &= \left[\frac{\tau_\varepsilon^{1/2}}{\varepsilon}\left(\check{\vartheta}_{\tau_\varepsilon} - \vartheta_0\right)\right]^2 \frac{\varepsilon}{\tau_\varepsilon} \to 0. \end{aligned}$$

Denote

$$\zeta^\star_{t,\varepsilon} = \frac{\vartheta^\star_{t,\varepsilon} - \vartheta_0}{\varepsilon}, \qquad \zeta^\star_t\left(\vartheta_0\right) = \mathbf{I}_0^t\left(\vartheta_0\right)^{-1} \int_0^t \frac{\dot S_*\left(\vartheta_0, \vartheta_0, s\right)}{\sigma\left(s\right)} \mathrm{d}w\left(s\right).$$

It is proved that for any $\tau_0 \in (0, T)$ the random process $\zeta^\star_{t,\varepsilon}, \tau_0 \le t \le T$ converges in distribution in $(\mathcal{C}\left[\tau_0, T\right], \mathfrak{B})$ to the Gaussian process $\zeta^\star_t\left(\vartheta_0\right), \tau_0 \le t \le T$. Recurrent presentation

$$\begin{aligned}\mathrm{d}\vartheta^\star_{t,\varepsilon} &= \mathbf{I}_0^t(\check\vartheta_{\tau_\varepsilon})^{-1} \frac{\dot S_*(\check\vartheta_{\tau_\varepsilon}, \vartheta_{\tau_\varepsilon}, s)^2}{\sigma\left(t\right)^2} \left[\check\vartheta_{\tau_\varepsilon} - \vartheta^\star_{t,\varepsilon}\right] \mathrm{d}t \\ &\quad + \mathbf{I}_0^t(\check\vartheta_{\tau_\varepsilon})^{-1} \frac{\dot S(\check\vartheta_{\tau_\varepsilon}, \check\vartheta_{\tau_\varepsilon}, t)}{\sigma\left(t\right)^2} \left[\mathrm{d}X_t - M(\check\vartheta_{\tau_\varepsilon}, t)\mathrm{d}t\right],\end{aligned}$$

Adaptive filter.

Suppose that we already have:

1. The learning interval $[0, \tau_\varepsilon]$, $\tau_\varepsilon = \varepsilon^{2/3}$.
2. Preliminary MDE $\check\vartheta_{\tau_\varepsilon}$.
3. K-B equations for $M(\check\vartheta_{\tau_\varepsilon}, s), \tau_\varepsilon \leq s \leq T$
4. One-step MLE-process $\vartheta^\star_{t,\varepsilon}, \tau_\varepsilon < t \leq T$:

$$\vartheta^\star_{t,\varepsilon} = \check\vartheta_{\tau_\varepsilon} + \mathbf{I}^t_{\tau_\varepsilon}(\check\vartheta_{\tau_\varepsilon})^{-1} \int_{\tau_\varepsilon}^t \frac{\dot S_*(\check\vartheta_{\tau_\varepsilon}, \check\vartheta_{\tau_\varepsilon}, s)}{\sigma\left(s\right)^2} \left[\mathrm{d}X_s - M(\check\vartheta_{\tau_\varepsilon}, s)\mathrm{d}s\right].$$

The adaptive K-B filter $m^\star_{t,\varepsilon}, \tau_\varepsilon \le s \le T$ is

$$\mathrm{d}m^\star_{t,\varepsilon} = \left[a\left(\vartheta^\star_{t,\varepsilon}, t\right) - \frac{\gamma_{t,\varepsilon}\ f\left(\vartheta^\star_{t,\varepsilon}, t\right)^2}{\sigma\left(t\right)^2} \right] m^\star_{t,\varepsilon}\mathrm{d}t + \frac{\gamma_{t,\varepsilon}\ f\left(\vartheta^\star_{t,\varepsilon}, t\right)}{\sigma\left(t\right)^2}\mathrm{d}X_t,$$

$$\frac{\partial \gamma_{t,\varepsilon}}{\partial t} = 2a\left(\vartheta^\star_{t,\varepsilon}, t\right)\gamma_{t,\varepsilon} - \frac{\gamma^2_{t,\varepsilon}\ f\left(\vartheta^\star_{t,\varepsilon}, t\right)^2}{\sigma\left(t\right)^2} + b\left(\vartheta^\star_{t,\varepsilon}, t\right)^2, \qquad \gamma_{0,\varepsilon} = 0.$$

The initial value $m^\star_{\tau_\varepsilon,\varepsilon} = m(\check{\vartheta}_{\tau_\varepsilon}, \tau_\varepsilon)$.

Introduce the notation:

$$A(\vartheta_0, t) = a(\vartheta_0, t) - \frac{f(\vartheta_0, t)^2 \gamma_*(\vartheta_0, t)}{\sigma(t)^2},$$

$$\Phi(\vartheta_0, s, t) = \exp\left(\int_s^t A(\vartheta_0, v)\,\mathrm{d}v\right), \quad B(\vartheta_0, t) = \frac{f(\vartheta_0, t)\gamma_*(\vartheta_0, t)}{\sigma(t)^2},$$

$$K(\vartheta_0, t) = \left[\dot a(\vartheta_0, t) - B(\vartheta_0, t)\dot f(\vartheta_0, t)\right] y_t(\vartheta_0),$$

$$\zeta_s(\vartheta_0) = \mathbf{I}_{\tau_*}^{s}(\vartheta_0)^{-1} \int_{\tau_*}^{s} \frac{\dot S_*(\vartheta_0, \vartheta_0, r)}{\sigma(r)} \mathrm{d}w(r),$$

$$n^\star(\vartheta_0, t) = \int_\tau^t \Phi(\vartheta_0, s, t) K(\vartheta_0, s) \zeta_s(\vartheta_0)\ \mathrm{d}s,$$

$$D^\star(\vartheta_0, t) = \mathbf{E}_{\vartheta_0}\left[n^\star(\vartheta_0, t)^2\right].$$

Theorem

Suppose that the regularity conditions hold. Then the error of approximation of $m(\vartheta_0, t)$ *by* $m^\star_{t,\varepsilon}$ *is asymptotically normal*

$$\frac{m^\star_{t,\varepsilon} - m(\vartheta_0, t)}{\varepsilon} \Longrightarrow n^\star(\vartheta_0, t), \qquad \tau \leq t \leq T$$

and for any $p \geq 2$

$$\varepsilon^{-p}\mathbf{E}_{\vartheta_0}\left|m^\star_{t,\varepsilon} - m(\vartheta_0, t)\right|^p \longrightarrow \mathbf{E}_{\vartheta_0}\left|n^\star(\vartheta_0, t)\right|^p.$$

If $p = 2$, *then* $\mathbf{E}_{\vartheta_0} n^\star(\vartheta_0, t)^2 = D^\star(\vartheta_0, t)$.

Asymptotic efficiency.

Introduce the function

$$\mathcal{S}_t^\star(\vartheta)^2 = \dot{y}_t(\vartheta,\vartheta)^2 \,\mathrm{I}_0^t(\vartheta)^{-1}, \qquad \vartheta \in \Theta.$$

For all estimators $\bar{m}_{t,\varepsilon}$ of $m(\vartheta_0, t)$ the following bound holds

$$\lim_{\nu\to 0} \underline{\lim}_{\varepsilon\to 0} \sup_{|\vartheta-\vartheta_0|\leq\nu} \varepsilon^{-2}\mathbf{E}_\vartheta \left|\bar{m}_{t,\varepsilon} - m(\vartheta, t)\right|^2 \geq \mathcal{S}_t^\star(\vartheta_0)^2,$$

The estimator $m_{t,\varepsilon}^*$ of $m(\vartheta_0, t)$ we call *asymptotically efficient*, if

$$\lim_{\nu\to 0} \lim_{\varepsilon\to 0} \sup_{|\vartheta-\vartheta_0|\leq\nu} \varepsilon^{-2}\mathbf{E}_\vartheta \left|m_{t,\varepsilon}^* - m(\vartheta, t)\right|^2 = \frac{\dot{y}_t(\vartheta_0,\vartheta_0)^2}{\mathrm{I}_0^t(\vartheta_0)}$$

holds for all $\vartheta_0 \in \Theta$.

Note that $m_{t,\varepsilon}^\star$ is not asymptotically efficient.

We have

$$m(\vartheta,t) = y_0\Phi(\vartheta,0,t) + B(\vartheta,t)X_t - \int_0^t X_s\Phi(\vartheta,s,t)\left[B(\vartheta,s)f(\vartheta,s) - a(\vartheta,s) + B_s'(\vartheta,s)\right]\mathrm{d}s.$$

Let us put $\Phi_\varepsilon(s,t) = \Phi\left(\vartheta^\star_{t,\varepsilon},s,t\right)$ and consider another adaptive filter

$$\hat m_{t,\varepsilon} = m\left(\vartheta^\star_{t,\varepsilon},t\right) = y_0\Phi_\varepsilon(0,t) + B\left(\vartheta^\star_{t,\varepsilon},t\right)X_t - \int_0^t X_s\Phi_\varepsilon(s,t)\left[B\left(\vartheta^\star_{t,\varepsilon},s\right)f\left(\vartheta^\star_{t,\varepsilon},s\right) - a\left(\vartheta^\star_{t,\varepsilon},s\right) + B_s'\left(\vartheta^\star_{t,\varepsilon},s\right)\right]\mathrm{d}s.$$

We prove that

$$\varepsilon^{-2}\mathbf{E}_\vartheta\left(\hat m_{t,\varepsilon} - m(\vartheta,t)\right)^2 \longrightarrow \dot y_t(\vartheta,\vartheta)^2\,\mathbf{E}_\vartheta\zeta_t(\vartheta)^2 = \mathcal{S}_t^*(\vartheta)^2.$$

Therefore $\hat m_{t,\varepsilon}$ is asymptotically efficient.

K.(1994) *Identification of Dynamical Systems with Small Noise.* Kluwer, Dordrecht.

K., Zhou, L. (2021) "On parameter estimation of the hidden Gaussian process in perturbed SDE", *Electr. J. of Stat.*, 15, 211-234

K. (2025) *Hidden Markov Processes and Adaptive Filtering*, Springer, Cham.

1. Adaptive filtering
2. Different levels of noise
3. The same level of noise ($\psi_\varepsilon = \varepsilon$)
4. Low noise in observations ($\psi_\varepsilon = 1$)
5. Different Levels ($\psi_\varepsilon = \varepsilon^\delta, 1/3 < \delta < 1$)
6. Different Levels ($\delta = 1/3$)
7. Different Levels ($0 < \delta < 1/3$)

Model of observations. Consider a non-homogeneous partially observed linear system described by the equations ($\delta = 0$)

$$\mathrm{d}X_t = f(\vartheta, t)\; Y_t\,\mathrm{d}t + \varepsilon\sigma(t)\,\mathrm{d}W_t, \qquad X_0 = 0,$$
$$\mathrm{d}Y_t = a(\vartheta, t)\, Y_t\,\mathrm{d}t + b(\vartheta, t)\,\mathrm{d}V_t, \qquad Y_0 = y_0.$$

Limit model ($\varepsilon = 0$) is

$$x_t' = f(\vartheta_0, t)\, Y_t, \qquad x_0 = 0, \qquad 0 \leq t \leq T,$$
$$\mathrm{d}Y_t = a(\vartheta_0, t)\, Y_t\mathrm{d}t + b(\vartheta_0, t)\,\mathrm{d}V_t, \qquad Y_0 = y_0.$$

Here ϑ can be estimated without error by "observations" x_t'.

The process $x'_t, 0 \leq t \leq T$ is described by the equation

$$\mathrm{d}x'_t = \left[f'\left(\vartheta_0, t\right) + f\left(\vartheta_0, t\right) a\left(\vartheta_0, t\right)\right] Y_t \mathrm{d}t + f\left(\vartheta_0, t\right) b\left(\vartheta_0, t\right) \mathrm{d}V_t,$$

where $x'_0 = 0$. Recall that by Itô's formula, we have

$$x'^2_t = 2\int_0^t x'_s \mathrm{d}x_s + \int_0^t b\left(\vartheta_0, s\right)^2 f\left(\vartheta_0, s\right)^2 \mathrm{d}s.$$

Thus the following function is deterministic

$$\Psi_t = x'^2_t - 2\int_0^t x'_s \mathrm{d}x'_s = \int_0^t b\left(\vartheta_0, s\right)^2 f\left(\vartheta_0, s\right)^2 \mathrm{d}s \equiv K\left(\vartheta_0, t\right).$$

Under mild identifiability conditions, the observed function Ψ_t defines ϑ_0 for any $t \in (0, T]$ without error. For example, the estimator ϑ^* defined by the equation $K(\vartheta^*, t) = \Psi_t$ is without error, i.e., $\vartheta^* = \vartheta_0$.

If $f(\vartheta, t) = f(t)$, $b(\vartheta, t)^2 = h(t) + \vartheta g(t)$, where all functions and ϑ are positive, then the "estimator"

$$\vartheta_t^* = \left(\int_0^t g(s) f(s)^2 \, \mathrm{d}s\right)^{-1} \left[\Psi_t - \int_0^t h(s) f(s)^2 \, \mathrm{d}s\right] = \vartheta_0.$$

Therefore the consistent estimation by observations X^T is not excluded.

Semiparametric estimation.

Suppose that we have the partially observed system

$$\mathrm{d}X_t = f(t)\, Y_t \mathrm{d}t + \varepsilon\sigma(t)\,\mathrm{d}W_t, \qquad X_0 = 0, \qquad 0 \leq t \leq T,$$
$$\mathrm{d}Y_t = a(t)\, Y_t \mathrm{d}t + b(t)\,\mathrm{d}V_t, \qquad Y_0 = y_0,$$

where $a(\cdot), b(\cdot), f(\cdot)$ and $\sigma(\cdot)$ are unknown continuously differentiable functions.

Consider the problem of estimation of the function

$$\Psi_\tau = \int_0^\tau f(t)^2\, b(t)^2\, \mathrm{d}t, \qquad 0 < \tau \leq T$$

by observations X^T. Remind that this is quadratic variation of the derivative of the process X_τ at the point $\varepsilon = 0$.

Introduce the statistic

$$\Psi_{\tau,\varepsilon} = \sum_{i=0}^{N_{\tau,\varepsilon}-1} \left(\frac{X_{t_{i+1}+\delta_\varepsilon} - X_{t_{i+1}}}{\delta_\varepsilon} - \frac{X_{t_i+\delta_\varepsilon} - X_{t_i}}{\delta_\varepsilon} \right)^2, \qquad 0 < \tau \leq T.$$

Here $t_i = i\phi_\varepsilon, N_{\tau,\varepsilon} = \left[\frac{\tau}{\phi_\varepsilon}\right]$, the rates $\varphi_\varepsilon \to 0, \delta_\varepsilon \to 0$ will be defind later. Just note that as the first step is derivation we wait that the rate $\delta_\varepsilon \to 0$ has to be faster than the step of discretization $\phi_\varepsilon \to 0$.

Proposition

Let the condition $a(\cdot), b(\cdot), f(\cdot), \sigma(\cdot) \in \mathcal{C}^1[0, T]$ hold and chose $\delta_\varepsilon = \varepsilon, \phi_\varepsilon = \varepsilon^{2/3}$. Then for any $p > 0$ there exists a constant $C > 0$ such that

$$\mathbf{E}\left|\Psi_{\tau,\varepsilon} - \Psi_\tau\right|^p \leq C\tau^{p/2}\varepsilon^{p/3}.$$

Parameter estimation

Remind the model of observations

$$\mathrm{d}X_t = f(\vartheta, t)\, Y_t\, \mathrm{d}t + \varepsilon\sigma(t)\, \mathrm{d}W_t, \qquad X_0 = 0, \qquad 0 \leq t \leq T,$$
$$\mathrm{d}Y_t = a(\vartheta, t)\, Y_t\, \mathrm{d}t + b(\vartheta, t)\, \mathrm{d}V_t, \qquad Y_0 = y_0$$

The functions $b(\cdot)$ and $f(\cdot)$ are supposed to be known. The functions $a(\cdot)$ and $\sigma(\cdot)$ do not used in the construction of the estimator. The parameter $\vartheta \in \Theta \subset \mathcal{R}^d$ is unknown and has to be estimated by observations X^T. The set Θ is open, convex and bounded.

Fix some $\tau \in (0, T]$ and define the function

$$\Psi_t(\vartheta) = \int_0^t f(\vartheta, s)^2 b(\vartheta, s)^2 \, \mathrm{d}s, \qquad 0 < t \leq \tau, \qquad \vartheta \in \Theta \subset \mathcal{R}^d.$$

As estimator of this function we take the nonparametric estimator

$$\Psi_{t,\varepsilon} = \sum_{i=0}^{N_{t,\varepsilon}-1} \left(\frac{X_{t_{i+1}+\varepsilon} - X_{t_{i+1}}}{\varepsilon} - \frac{X_{t_i+\varepsilon} - X_{t_i}}{\varepsilon} \right)^2, \qquad 0 < t \leq \tau.$$

Here $t_{i+1} - t_i = \varepsilon^{1/3}$, $t_0 = 0$ and $N_{t,\varepsilon} = \left[t\varepsilon^{-1/3} \right]$.

We already know that

$$\Psi_{t,\varepsilon} \longrightarrow \Psi_t(\vartheta_0).$$

This convergence allow us the following constructions of the following estimators. .

MME. The MME $\vartheta^*_{\tau,\varepsilon}$ can be defined as follows. Let us fix d points $0 < \tau_1 < \tau_2 < \ldots < \tau_d = \tau$ and denote the vectors $\Psi^*_{\tau,\varepsilon} = \left(\Psi_{\tau_1,\varepsilon}, \Psi_{\tau_2,\varepsilon}, \ldots, \Psi_{\tau_d,\varepsilon}\right)^\top$ and $\Psi^*_\tau\left(\vartheta\right) = \left(\Psi_{\tau_1}\left(\vartheta\right), \Psi_{\tau_2}\left(\vartheta\right), \ldots, \Psi_{\tau_d}\left(\vartheta\right)\right)^\top$. The estimator is defined as solution of the system of equations

$$\Psi^*_{\tau,\varepsilon} = \Psi^*_\tau(\vartheta^*_{\tau,\varepsilon}).$$

Note that having the estimator $\Psi_{t,\varepsilon}, 0 \leq t \leq \tau$ we can define the MDE $\check{\vartheta}_{\tau,\varepsilon}$ too by the equation

$$\int_0^\tau \left[\Psi_{t,\varepsilon} - \Psi_t(\check{\vartheta}_{\tau,\varepsilon})\right]^2 \mathrm{d}t = \inf_{\vartheta\in\Theta} \int_0^\tau \left[\Psi_{t,\varepsilon} - \Psi_t\left(\vartheta\right)\right]^2 \mathrm{d}t.$$

Introduce the notation

$$g\left(\vartheta_0,\nu\right) = \inf_{\|\vartheta-\vartheta_0\|\geq\nu} \left\|\Psi_\tau\left(\vartheta\right) - \Psi_\tau\left(\vartheta_0\right)\right\|, \quad \mathbf{T}\left(\vartheta_0\right) = \dot\Psi_\tau^*(\vartheta_0)\dot\Psi_\tau^*(\vartheta_0)^\top$$

and *Conditions* $\mathscr{H}^*$:

$\mathscr{H}_1^*$. *The functions* $a\left(\vartheta,t\right), b\left(\vartheta,t\right), f\left(\vartheta,t\right), \sigma\left(t\right), \in \mathcal{C}_t^1, \mathcal{C}_\vartheta^2$.

$\mathscr{H}_2^*$. *The matrix* $\mathbf{T}\left(\vartheta\right)$ *is uniformly non degenerate*

$$\inf_{\vartheta\in\Theta} \inf_{\|\mathrm{e}\|=1,\mathrm{e}\in\mathcal{R}^d} \mathrm{e}^\top \mathbf{T}\left(\vartheta\right)\mathrm{e} > 0.$$

$\mathscr{H}_3^*$. *For any* $\nu > 0$

$$\inf_{\vartheta_0\in\Theta} g\left(\vartheta_0,\nu\right) > 0.$$

Proposition

Let the conditions $\mathscr{H}^*$ hold. Then the MME $\vartheta^*_{\tau,\varepsilon}$ is uniformly consistent and for any $p \geq 2$ there exists a constant $C = C(p) > 0$ such that

$$\sup_{\vartheta_0 \in \Theta} \varepsilon^{-p/3}\, \tau^{p/2} \mathbf{E}_{\vartheta_0} \left\| \vartheta^*_{\tau,\varepsilon} - \vartheta_0 \right\|^p \leq C.$$

Example 1.

Suppose that $f(\vartheta, t) = \vartheta f(t), \vartheta \in (\alpha, \beta), \alpha > 0,\ b(\vartheta, t) = b(t)$.

$$\Psi_\tau(\vartheta) = \vartheta^2 \int_0^\tau f(t)^2\, b(t)^2\, \mathrm{d}t$$

and the MME is

$$\vartheta^*_{\tau,\varepsilon} = \sqrt{\Psi_{\tau,\varepsilon}} \left(\int_0^\tau f(t)^2\, b(t)^2\, \mathrm{d}t \right)^{-1/2}.$$

Example 2.

Let $f(\vartheta, t) = f(t)$ and $b(\vartheta, t) = \Big(h(t) + \sum_{k=1}^{d} \vartheta_k g_k(t)\Big)^{1/2}$. Then

$$\Psi_{\tau_k}(\vartheta) = \int_0^{\tau_k} f(s)^2\, h(s)\,\mathrm{d}s + \sum_{j=1}^{d} \vartheta_j \int_0^{\tau_k} f(s)^2\, g_j(s)\,\mathrm{d}s$$

$$= H_k + \sum_{j=1}^{d} \vartheta_j G_{jk}, \qquad\qquad k = 1, \ldots, d.$$

Suppose that the matrix $\mathbf{G} = (G_{jk})_{d\times d}$ is non degenerate. Then we have the relation

$$\vartheta = \mathbf{G}^{-1}\Big(\Psi_\tau(\vartheta) - \mathrm{H}\Big), \qquad\qquad \mathrm{H} = (H_1, \ldots, H_d)^\top$$

with the obvious notations.

Hence the MME is $\vartheta^*_{\tau,\varepsilon} = \mathbf{G}^{-1}\Big(\Psi_{\tau,\varepsilon} - \mathrm{H}\Big)$ and this estimator has the properties described in the Proposition.

MLE and BE.

The LR function is

$$L(\vartheta, X^T) = \exp\left\{\int_0^T \frac{M(\vartheta,t)}{\varepsilon^2\sigma(t)^2}\,\mathrm{d}X_t - \int_0^T \frac{M(\vartheta,t)^2}{2\varepsilon^2\sigma(t)^2}\,\mathrm{d}t\right\}, \quad \vartheta \in \Theta.$$

where $M(\vartheta,t) = f(\vartheta,t)\,m(\vartheta,t)$ and

$$\mathrm{d}m(\vartheta,t) = \left[a(\vartheta,t) - \frac{\gamma(\vartheta,t)\,f(\vartheta,t)^2}{\varepsilon^2\sigma(t)^2}\right] m(\vartheta,t)\,\mathrm{d}t + \frac{\gamma(\vartheta,t)\,f(\vartheta,t)}{\varepsilon^2\sigma(t)^2}\mathrm{d}X_t,$$

$$\frac{\partial\gamma(\vartheta,t)}{\partial t} = 2a(\vartheta,t)\,\gamma(\vartheta,t) - \frac{\gamma(\vartheta,t)^2 f(\vartheta,t)^2}{\varepsilon^2\sigma(t)^2} + b(\vartheta,t)^2,$$

Then the MLE $\hat\vartheta_\varepsilon$ and BE $\tilde\vartheta_\varepsilon$ are defined by the usual relations

$$L(\hat\vartheta_\varepsilon, X^T) = \sup_{\vartheta\in\Theta} L(\vartheta, X^T), \qquad \tilde\vartheta_\varepsilon = \frac{\int_\Theta \vartheta\, p(\vartheta) L(\vartheta, X^T)\,\mathrm{d}\vartheta}{\int_\Theta p(\vartheta) L(\vartheta, X^T)\,\mathrm{d}\vartheta}.$$

Denote $S(\vartheta, t) = f(\vartheta, t)\, b(\vartheta, t)$ and set

$$\begin{aligned} G(\vartheta, \vartheta_0) &= \int_0^T \frac{\left[f(\vartheta, t)\, b(\vartheta, t) - f(\vartheta_0, t)\, b(\vartheta_0, t)\right]^2}{2b(\vartheta, t)\, f(\vartheta, t)\, \sigma(t)} \mathrm{d}t \\ &= \int_0^T \frac{\left[S(\vartheta, t) - S(\vartheta_0, t)\right]^2}{2S(\vartheta, t)\, \sigma(t)} \mathrm{d}t, \\ \mathbf{I}(\vartheta) &= \int_0^T \frac{f(\vartheta, t)\, b(\vartheta, t)}{2\sigma(t)} \left(\frac{\dot f(\vartheta, t)}{f(\vartheta, t)} + \frac{\dot b(\vartheta, t)}{b(\vartheta, t)}\right)^2 \mathrm{d}t \\ &= \int_0^T \frac{S(\vartheta, t)}{2\sigma(t)} \left[\frac{\partial}{\partial\vartheta} \ln S(\vartheta, t)\right]^2 \mathrm{d}t. \end{aligned}$$

Conditions $\hat{\mathscr{H}}$:

$\hat{\mathscr{H}}_1$. $f(\vartheta,\cdot), a(\vartheta,\cdot), b(\vartheta,\cdot) \in \mathcal{C}_\vartheta^{2,b}$.

$\hat{\mathscr{H}}_2$. $f(\vartheta,t), \sigma(t)$, $b(\vartheta,t)$ are separated from zero.

$\hat{\mathscr{H}}_3$. For any $\nu > 0$,

$$\inf_{\vartheta_0\in\Theta} \inf_{\|\vartheta-\vartheta_0\|>\nu} G(\vartheta,\vartheta_0) > 0.$$

$\hat{\mathscr{H}}_4$. The Fisher information matrix is uniformly non degenerate.

Theorem

Suppose that the conditions $\hat{\mathscr{H}}$ hold. Then the MLE $\hat\vartheta_\varepsilon$ and BE $\tilde\vartheta_\varepsilon$ are uniformly on compacts $\mathbb{K} \subset \Theta$ consistent, asymptotically normal, i.e.,

$$\frac{\left(\hat\vartheta_\varepsilon - \vartheta_0\right)}{\sqrt{\varepsilon}} \Longrightarrow \zeta\left(\vartheta_0\right) \sim \mathcal{N}\left(0, \mathbf{I}\left(\vartheta_0\right)^{-1}\right), \quad \frac{\left(\tilde\vartheta_\varepsilon - \vartheta_0\right)}{\sqrt{\varepsilon}} \Longrightarrow \zeta\left(\vartheta_0\right),$$

the moments converge: for any $p \geq 0$ uniformly on $\mathbb{K} \subset \Theta$

$$\lim_{\varepsilon \to 0} \varepsilon^{p/2} \mathbf{E}_{\vartheta_0} \left\| \left(\hat\vartheta_\varepsilon - \vartheta_0\right)\right\|^p = \mathbf{E}_{\vartheta_0} \left\|\zeta\left(\vartheta_0\right)\right\|^p,$$

$$\lim_{\varepsilon \to 0} \varepsilon^{p/2} \mathbf{E}_{\vartheta_0} \left\| \left(\tilde\vartheta_\varepsilon - \vartheta_0\right)\right\|^p = \mathbf{E}_{\vartheta_0} \left\|\zeta\left(\vartheta_0\right)\right\|^p$$

and both estimators are asymptotically efficient.

Lemma

Suppose that the functions $F(t)$, $G(t)$, $H(t)$, $t \in [0, T]$ are continuously differentiable on $[0, T]$, that $F(0) = 0$ and that $F(t) > 0$ for all $t \in (0, T]$. Then, for any $t \in (0, T]$, we have, as $\varepsilon \to 0$, the relations

$$N_\varepsilon(t) = \frac{1}{\varepsilon}\int_0^t \exp\left\{-\frac{1}{\varepsilon}\int_s^t F(v)\,\mathrm{d}v\right\} G(s)\,\mathrm{d}s = \frac{G(t)}{F(t)}\{1 + O(\varepsilon)\}$$

$$M_\varepsilon(t) = \frac{1}{\sqrt{\varepsilon}}\int_0^t \exp\left\{-\frac{1}{\varepsilon}\int_s^t F(v)\,\mathrm{d}v\right\} H(s)\,\mathrm{d}W_s \Longrightarrow \frac{H(t)}{\sqrt{2F(t)}}\,\xi_t.$$

Here, the variables $\xi_t \sim \mathcal{N}(0, 1)$ are mutually independent for all $t \in (0, T]$.

The family $(\xi_t, t \in (0, T))$ is not a stochastic process!

Lemma

For any $t_0 \in (0, T]$, *we have*

$$\sup_{t_0 \leq t \leq T} \left| \frac{\gamma(\vartheta, t)}{\varepsilon} - \frac{b(\vartheta, t)\,\sigma(t)}{f(\vartheta, t)} \right| = O(\varepsilon).$$

The equation for the filter are replaced by the equation

$$\begin{aligned} \mathrm{d}m(\vartheta, t) &= \left[a(\vartheta, t) - \frac{\gamma_*(\vartheta, t)\, f(\vartheta, t)^2}{\varepsilon\sigma(t)^2} \right] m(\vartheta, t)\,\mathrm{d}t \\ &\qquad + \frac{\gamma_*(\vartheta, t)\, f(\vartheta, t)}{\varepsilon\sigma(t)^2}\mathrm{d}X_t \\ &= A(\vartheta, t)\, m(\vartheta, t)\,\mathrm{d}t + B(\vartheta, t)\,\mathrm{d}X_t, \qquad m(\vartheta, 0) = y_0. \end{aligned}$$

Here

$$\gamma_*(\vartheta, t) = \frac{b(\vartheta, t)\,\sigma(t)}{f(\vartheta, t)}.$$

One-step MLE-process.

Introduce notations: $M\left(\vartheta^*_{\tau,\varepsilon},t\right)=f\left(\vartheta^*_{\tau,\varepsilon},t\right)\hat m\left(t\right),$

$$\dot M\left(\vartheta^*_{\tau,\varepsilon},t\right)=\dot f\left(\vartheta^*_{\tau,\varepsilon},t\right)\hat m\left(t\right)+f\left(\vartheta^*_{\tau,\varepsilon},t\right)\hat{\dot m}\left(t\right),$$

$$\mathbf{I}_\tau^t\left(\vartheta\right)=\int_\tau^t\frac{\dot S\left(\vartheta,s\right)^2}{2S\left(\vartheta,s\right)\sigma\left(s\right)}{\rm d}s,\qquad\qquad \zeta_{t,\varepsilon}\left(\vartheta_0\right)=\frac{\vartheta^\star_{t,\varepsilon}-\vartheta_0}{\sqrt{\varepsilon}},$$

$$\zeta_t\left(\vartheta_0\right)=\mathbf{I}_\tau^t(\vartheta_0)^{-1}\int_\tau^t\frac{\dot S\left(\vartheta_0,s\right)}{\sqrt{2S\left(\vartheta_0,s\right)\sigma\left(s\right)}}\,{\rm d}w\left(s\right),\qquad\qquad \tau<t\leq T,$$

where $w\left(s\right),\tau\leq s\leq T$ is some Wiener process and

$$\begin{aligned}{\rm d}\hat m\left(t\right)&=a\left(\vartheta^*_{\tau,\varepsilon},t\right)\hat m\left(t\right){\rm d}t\\&\qquad+\frac{\gamma\left(\vartheta^*_{\tau,\varepsilon},t\right)f\left(\vartheta^*_{\tau,\varepsilon},t\right)}{\varepsilon^2\sigma\left(t\right)^2}\left[{\rm d}X_t-f\left(\vartheta^*_{\tau,\varepsilon},t\right)\hat m\left(t\right){\rm d}t\right],\end{aligned}$$

$${\rm d}\hat{\dot{\rm m}}\left(t\right)=A\left(\vartheta^*_{\tau,\varepsilon},t\right)\hat{\dot{\rm m}}\left(t\right){\rm d}t+\dot{\rm A}\left(\vartheta^*_{\tau,\varepsilon},t\right)\hat m\left(t\right){\rm d}t+\dot{\rm B}\left(\vartheta^*_{\tau,\varepsilon},t\right){\rm d}X_t.$$

Introduce the One-step MLE-process $\vartheta^\star_{t,\varepsilon}, \tau < t \leq T$

$$\vartheta^\star_{t,\varepsilon} = \vartheta^*_{\tau,\varepsilon} + \mathbf{I}_\tau^t(\vartheta^*_{\tau,\varepsilon})^{-1} \int_\tau^t \frac{\dot{\mathrm{M}}(\vartheta^*_{\tau,\varepsilon}, s)}{\varepsilon\sigma\left(s\right)^2}\left[\mathrm{d}X_s - M(\vartheta^*_{\tau,\varepsilon}, s)\mathrm{d}s\right],$$

and *Conditions* $\mathscr{H}^\star$:

$\mathscr{H}_1^\star$. *The functions* $f\left(\vartheta, t\right), b\left(\vartheta, t\right), \sigma\left(t\right); \left(\vartheta, t\right) \in \bar{\Theta} \times [0, T]$ *are separated from zero.*

$\mathscr{H}_2^\star$. *The conditions* $\mathscr{H}^*$ *of Proposition 4 hold.*

$\mathscr{H}_3^\star$. *For any* $t_0 \in (\tau, T]$

$$\inf_{\vartheta_0\in\Theta} \quad \inf_{\|\mathrm{e}\|=1,\mathrm{e}\in\mathcal{R}^d} \mathrm{e}^\top \mathbf{I}_\tau^{t_0}\left(\vartheta_0\right)\mathrm{e} > 0.$$

Theorem

Let the assumptions $\mathscr{H}^\star$ hold. Then the One-step MLE-process $\vartheta^\star_{t,\varepsilon}$, $\tau < t \leq T$ has the following properties:

1. *It is uniformly consistent: for any $t_0 \in (\tau, T]$ and any $\nu > 0$*

$$\mathbf{P}_{\vartheta_0}\left(\sup_{t_0\leq t\leq T}\left\|\vartheta^\star_{t,\varepsilon} - \vartheta_0\right\| \geq \nu\right) \longrightarrow 0,$$

2. *The random process $\zeta_{t,\varepsilon}(\vartheta_0)$, $t_0 \leq t \leq T$ converges in distribution in the measurable space $(\mathcal{C}[t_0, T], \mathfrak{B}])$ to the Gaussian process $\zeta_t(\vartheta_0)$, $t_0 \leq t \leq T$*
3. *All polynomial moments converge: for any $p > 0$*

$$\varepsilon^{-p/2}\mathbf{E}_{\vartheta_0}\|\vartheta^\star_{t,\varepsilon} - \vartheta_0\|^p \longrightarrow \mathbf{E}_{\vartheta_0}\|\zeta_t(\vartheta_0)\|^p.$$

Adaptive filter.

The adaptive filter is constructed as follows.

- *Fix the learning interval* $[0, \tau_\varepsilon]$*, where* $\tau_\varepsilon = \varepsilon^{1/12}$ *with the observations* $X^{\tau_\varepsilon} = (X_t, 0 \leq t \leq \tau_\varepsilon)$*.*
- *Put* $t_0 = 0, t_{i+1} = t_i + \varepsilon^{1/3}$*,* $N_{\tau_\varepsilon,\varepsilon} = \left[\varepsilon^{-1/4}\right]$ *and calculate the statistic* $\hat{\Psi}_{\tau_\varepsilon,\varepsilon}$*.*
- *Solve the equation* $\Psi_{\tau_\varepsilon}(\vartheta) = \hat{\Psi}_{\tau_\varepsilon,\varepsilon}$ *whose solution is* $\vartheta = \vartheta^*_{\tau_\varepsilon}$*.*
- *Introduce the One-step MLE-process* $\vartheta^\star_{t,\varepsilon}, \tau_\varepsilon < t \leq T$,

$$\vartheta^\star_{t,\varepsilon} = \vartheta^*_{\tau_\varepsilon} + \frac{1}{\mathrm{I}^t_{\tau_\varepsilon}\left(\vartheta^*_{\tau_\varepsilon}\right)} \int_{\tau_\varepsilon}^{t} \frac{\dot{M}\left(\vartheta^*_{\tau_\varepsilon}, s\right)}{\varepsilon\sigma\left(s\right)^2} \left[\mathrm{d}X_s - M\left(\vartheta^*_{\tau_\varepsilon}, s\right)\mathrm{d}s\right],$$

Here $M\left(\vartheta^*_{\tau_\varepsilon}, t\right) = f\left(\vartheta^*_{\tau_\varepsilon}, t\right) m\left(\vartheta^*_{\tau_\varepsilon}, t\right), \tau_\varepsilon < t \leq T$

- $M\left(\vartheta^*_{\tau_\varepsilon}, t\right), \tau_\varepsilon < t \leq T$ *is solution of the equation*

$$\mathrm{d}M\left(\vartheta^*_{\tau_\varepsilon}, t\right) = -\frac{b\left(\vartheta^*_{\tau_\varepsilon}, t\right) f\left(\vartheta^*_{\tau_\varepsilon}, t\right)}{\varepsilon\sigma\left(t\right)} M\left(\vartheta^*_{\tau_\varepsilon}, t\right) \mathrm{d}t + \frac{b\left(\vartheta^*_{\tau_\varepsilon}, t\right) f\left(\vartheta^*_{\tau_\varepsilon}, t\right)}{\varepsilon\sigma\left(t\right)} \mathrm{d}X_t,$$

- $\dot M\left(\vartheta^*_{\tau_\varepsilon}, t\right), \tau_\varepsilon < t \leq T$ *is solution of the equation*

$$\mathrm{d}\dot M\left(\vartheta^*_{\tau_\varepsilon}, t\right) = -\frac{b\left(\vartheta^*_{\tau_\varepsilon}, t\right) f\left(\vartheta^*_{\tau_\varepsilon}, t\right)}{\varepsilon\sigma\left(t\right)} \dot M\left(\vartheta^*_{\tau_\varepsilon}, t\right) \mathrm{d}t + \frac{\dot b\left(\vartheta^*_{\tau_\varepsilon}, t\right) f\left(\vartheta^*_{\tau_\varepsilon}, t\right) + \dot f\left(\vartheta^*_{\tau_\varepsilon}, t\right) b\left(\vartheta, t\right)}{\varepsilon\sigma\left(t\right)} \left[\mathrm{d}X_t - M\left(\vartheta^*_{\tau_\varepsilon}, t\right) \mathrm{d}t\right].$$

The *adaptive filter* is given by the equation

$$\mathrm{d}m^\star_{t,\varepsilon} = -\frac{b\left(\vartheta^\star_{t,\varepsilon}, t\right) f\left(\vartheta^\star_{t,\varepsilon}, t\right)}{\varepsilon\sigma\left(t\right)} m^\star_{t,\varepsilon}\mathrm{d}t + \frac{b\left(\vartheta^\star_{t,\varepsilon}, t\right)}{\varepsilon\sigma\left(t\right)}\mathrm{d}X_t, \quad \tau_\varepsilon \leq t \leq T$$

subject to the initial value $m^\star_{\tau_\varepsilon,\varepsilon} = m(\check{\vartheta}_{\tau_\varepsilon}, \tau_\varepsilon)$. Note that in this case too we have the recursive representation

$$\begin{aligned}\mathrm{d}\vartheta^\star_{t,\varepsilon} &= \frac{\left[\vartheta^*_{\tau_\varepsilon} - \vartheta^\star_{t,\varepsilon}\right]}{\mathrm{I}^t_{\tau_\varepsilon}\left(\vartheta^*_{\tau_\varepsilon}\right)} \frac{\dot{S}\left(\vartheta^*_{\tau_\varepsilon}, t\right)^2}{2S\left(\vartheta^*_{\tau_\varepsilon}, t\right)\sigma\left(t\right)} \\ &\quad + \frac{1}{\mathrm{I}^t_{\tau_\varepsilon}\left(\vartheta^*_{\tau_\varepsilon}\right)} \frac{\dot{M}\left(\vartheta^*_{\tau_\varepsilon}, t\right)}{\varepsilon\sigma\left(t\right)^2}\left[\mathrm{d}X_t - M\left(\vartheta^*_{\tau_\varepsilon}, s\right)\mathrm{d}t\right],\end{aligned}$$

Theorem

Let the regularity conditions hold. Then,

1. *If ϑ_0 is known then*

$$\frac{Y_t - m(\vartheta_0, t)}{\sqrt{\varepsilon}} = \sqrt{\frac{b(\vartheta_0, t)\,\sigma(t)}{2f(\vartheta_0, t)}}\,\left(\xi_{t,\varepsilon}^{(1)} - \xi_{t,\varepsilon}^{(2)}\right)\left(1 + O\left(\varepsilon^{1/2}\right)\right).$$

2. *If ϑ_0 is unknown, then we have the representation*

$$\frac{Y_t - m_t^\star}{\sqrt{\varepsilon}} = \frac{\dot f(\vartheta_0, t)}{f(\vartheta_0, t)}\hat\zeta_{t,\varepsilon}(\vartheta_0)\;Y_t + \sqrt{\frac{b(\vartheta_0, t)\,\sigma(t)}{2f(\vartheta_0, t)}}\left(\xi_{t,\varepsilon}^{(1)} - \xi_{t,\varepsilon}^{(2)}\right) +$$

where $q_* > 0$ and

$$\mathbf{E}_{\vartheta_0}\left(\hat\zeta_{t,\varepsilon}(\vartheta_0)\;Y_t\right)^2 \longrightarrow \mathbf{E}_{\vartheta_0}\hat\zeta_t(\vartheta_0)^2\;\mathbf{E}_{\vartheta_0}Y_t^2 = \mathrm{I}_0^t(\vartheta_0)^{-1}\;\mathbf{E}_{\vartheta_0}Y_t^2.$$

Let us denote

$$\mathcal{S}_t^\star(\vartheta)^2 = \frac{\dot{f}(\vartheta, t)^2 \ \mathbf{E}_\vartheta Y_t^2}{f(\vartheta, t)^2 \, \mathrm{I}_0^t(\vartheta)}$$

where

$$\mathbf{E}_\vartheta Y_t^2 = y_0^2 \phi(\vartheta, 0, t)^2 + \int_0^t \phi(\vartheta, s, t)^2 \, b(\vartheta, s)^2 \, \mathrm{d}s.$$

We have the lower minimax bound: for any $t \in (0, T]$ and any $\vartheta_0 \in \Theta$

$$\lim_{\nu \to 0} \varliminf_{\varepsilon \to 0} \sup_{|\vartheta - \vartheta_0| \leq \nu} \varepsilon^{-1} \mathbf{E}_\vartheta \left| \bar{m}(t) - m(\vartheta, t) \right|^2 \geq \mathcal{S}_t^\star(\vartheta_0)^2 .$$

The adaptive filter $m_t^\star$ is asymptotically efficient

$$\lim_{\nu \to 0} \lim_{\varepsilon \to 0} \sup_{|\vartheta - \vartheta_0| \leq \nu} \varepsilon^{-1} \mathbf{E}_\vartheta \left| m_t^\star - m(\vartheta, t) \right|^2 = \mathcal{S}_t^\star(\vartheta_0)^2 .$$

- K (2019) "On parameter estimation of hidden Ornstein -Uhlenbeck process", *J. Multivariate Analysis*. 169, 1, 248-263.
- K (2024) "Volatility estimation of hidden Markov process and adaptive filtration", *Stoch. Processes Appl.*, July, 173, 104381

1 Adaptive filtering

2 Different levels of noise

3 The same level of noise ($\psi_\varepsilon = \varepsilon$)

4 Low noise in observations ($\psi_\varepsilon = 1$)

5 Different Levels ($\psi_\varepsilon = \varepsilon^\delta, 1/3 < \delta < 1$)

6 Different Levels ($\delta = 1/3$)

7 Different Levels ($0 < \delta < 1/3$)

Model of observations.

Consider the partially observed stochastic linear system

$$\begin{aligned} \mathrm{d}X_t &= f(\vartheta,t)\,Y_t\,\mathrm{d}t + \varepsilon\sigma(t)\,\mathrm{d}W_t, \qquad X_0 = 0, \qquad 0 \le t \le T,\\ \mathrm{d}Y_t &= a(\vartheta,t)\,Y_t\,\mathrm{d}t + \psi_\varepsilon b(\vartheta,t)\,\mathrm{d}V_t, \quad Y_0 = y_0 \ne 0, \end{aligned}$$

where $\psi_\varepsilon = \varepsilon^\delta, \delta \in (1/3, 1)$. The Kalman-Bucy filter is

$$\begin{aligned} \mathrm{d}m(\vartheta,t) &= a(\vartheta,t)m(\vartheta,t)\,\mathrm{d}t\\ &\quad + \frac{f(\vartheta,t)\,\gamma(\vartheta,t)}{\varepsilon^2\sigma(t)^2}\left[\mathrm{d}X_t - f(\vartheta,t)\,m(\vartheta,t)\,\mathrm{d}t\right], \quad m(\vartheta,0) = y_0. \end{aligned}$$

Riccati equation

$$\frac{\partial \gamma(\vartheta, t)}{\partial t} = 2a(\vartheta, t)\gamma(\vartheta, t) - \frac{f(\vartheta, t)^2 \gamma(\vartheta, t)^2}{\varepsilon^2 \sigma(t)^2} + \psi_\varepsilon^2 b(\vartheta, t)^2,$$

where $\gamma(\vartheta, 0) = 0$.

The likelihood ratio function is

$$L(\vartheta, X^T) = \exp\left\{\int_0^T \frac{M(\vartheta, t)}{\varepsilon^2 \sigma(t)^2} \mathrm{d}X_t - \int_0^T \frac{M(\vartheta, t)^2}{2\varepsilon^2 \sigma(t)^2} \mathrm{d}t\right\}, \qquad \vartheta \in \Theta.$$

Here $M(\vartheta, t) = f(\vartheta, t)\, m(\vartheta, t)$.

Parameter estimation

MDE. Fixed observation interval.

The MDE $\check\vartheta_{\tau,\varepsilon}$ is defined by the equation

$$\int_0^\tau \left[X_t - x_t(\check\vartheta_\varepsilon)\right]^2 \mathrm{d}t = \inf_{\vartheta\in\Theta} \int_0^\tau \left[X_t - x_t\left(\vartheta\right)\right]^2 \mathrm{d}t.$$

Introduce the $d \times d$ matrix

$$\check{\mathbf{T}}\left(\vartheta\right) = \int_0^\tau \dot{\mathrm{x}}_t\left(\vartheta\right) \dot{\mathrm{x}}_t\left(\vartheta\right)^\top \mathrm{d}t, \qquad \vartheta \in \Theta,$$

Gaussian vector

$$\check\zeta\left(\vartheta\right) = \check{\mathbf{T}}\left(\vartheta\right)^{-1} \int_0^\tau \int_0^t b\left(\vartheta, r\right) \int_r^t f\left(\vartheta, s\right) \phi\left(\vartheta, r, s\right) \mathrm{d}s\, \mathrm{d}V_r\, \dot{\mathrm{x}}_t\left(\vartheta\right)\, \mathrm{d}t,$$

where $\phi\left(\vartheta, r, s\right) = \exp\left(\int_r^s a\left(\vartheta, v\right) \mathrm{d}v\right)$ and the function

$$\check g\left(\vartheta_0, \nu\right) \equiv \inf_{\left|\vartheta-\vartheta_0\right|>\nu} \int_0^\tau \left[x_t\left(\vartheta_0\right) - x_t\left(\vartheta\right)\right]^2 \mathrm{d}t.$$

Conditions $\check{\mathscr{Z}}$.

$\check{\mathscr{Z}}_0$. *The set* $\Theta \subset \mathcal{R}^d$ *is open, convex and bounded.*

$\check{\mathscr{Z}}_1$. *The functions* $f(\vartheta, t), a(\vartheta, t), b(\vartheta, t)$, $0 \leq t \leq \tau, \vartheta \in \Theta$ *have two continuous bounded derivatives on* $\vartheta \in \bar{\Theta}$.

$\check{\mathscr{Z}}_2$. *The matrix* $\check{\mathbf{T}}(\vartheta)$ *is uniformly non degenerate*

$$\inf_{\vartheta\in\Theta} \inf_{\|\mathrm{e}\|=1, \mathrm{e}\in\mathcal{R}^d} \mathrm{e}\, \check{\mathbf{T}}(\vartheta)\, \mathrm{e}^\top > 0.$$

$\check{\mathscr{Z}}_2$. *For any* $\nu > 0$

$$\inf_{\vartheta_0\in\Theta} \check{g}(\vartheta_0, \nu) > 0.$$

Theorem

Let the conditions $\check{\mathscr{L}}$ hold. Then the MDE is consistent, asymptotically normal

$$\psi_\varepsilon^{-1}\left(\check{\vartheta}_{\tau,\varepsilon} - \vartheta_0\right) \Longrightarrow \check{\zeta}\left(\vartheta_0\right)$$

and the polynomial moments converge: for any $p > 0$

$$\psi_\varepsilon^{-p}\mathbf{E}_{\vartheta_0}\|\check{\vartheta}_{\tau,\varepsilon} - \vartheta_0\|^p \longrightarrow \mathbf{E}_{\vartheta_0}\|\check{\zeta}\left(\vartheta_0\right)\|^p.$$

MDE. Vanishing observation interval.

Let us describe the properties of MDE

$$\check{\vartheta}_{\tau_\varepsilon} = \arg \inf_{\vartheta \in \Theta} \int_0^{\tau_\varepsilon} \left[X_t - x_t\left(\vartheta\right)\right]^2 \mathrm{d}t$$

constructed by the observations $X^{\tau_\varepsilon} = (X_t, 0 \leq t \leq \tau_\varepsilon)$ on the vanishing interval $[0, \tau_\varepsilon]$, where $\tau_\varepsilon \to 0$. We will need such estimator for the construction of One-step MLE-process later.

Conditions $\mathscr{L}^\circ$.

$\mathscr{L}_0^\circ$. *The set* $\Theta = (\alpha, \beta)$, *where* α *and* β *are finite.*

$\mathscr{L}_1^\circ$. *The functions* $f(\vartheta, t), a(\vartheta, t), \vartheta \in \Theta, t \in [0, t_0]$ *have two continuous bounded derivatives on* ϑ. *Here* $t_0 > 0$ *is some (small) real from the interval* $[0, T]$.

$\mathscr{L}_2^\circ$. *The functions* $f(\vartheta, 0), \dot f(\vartheta, 0)$ *and* $b(\vartheta, 0)$ *are separated from zero uniformly on* $\vartheta \in \Theta$.

$\mathscr{L}_3^\circ$. *The functions* $f(\vartheta, t), a(\vartheta, t), b(\vartheta, t), \vartheta \in \Theta, t \in [0, t_0]$ *have continuous bounded derivatives on* t.

Let us denote

$$D(\vartheta) = \frac{1}{y_0^2 \dot{f}(\vartheta,0)^2} \left(\frac{33}{140} f(\vartheta,0)^2 b(\vartheta,0)^2 + \frac{6}{5} \sigma(0)^2 \right).$$

Proposition

Let conditions $\mathscr{L}^\circ$ hold and $\sqrt{\tau_\varepsilon}\psi_\varepsilon^{-1} \to \infty$. Then the MDE $\check{\vartheta}_{\tau_\varepsilon}$ is consistent and asymptotically normal

$$\sqrt{\frac{\psi_\varepsilon}{\varepsilon}} \left(\check{\vartheta}_{\tau_\varepsilon} - \vartheta_0 \right) \Longrightarrow \mathcal{N}\left(0, D(\vartheta_0)\right).$$

MLE and BE

Denote $\ell(\vartheta, t) = \partial \ln f(\vartheta, t) / \partial t = f'_t(\vartheta, t) / f(\vartheta, t)$. Then the Fisher information in this problem can be written as follows

$$\mathrm{I}(\vartheta) = \int_0^T \frac{[\dot a(\vartheta, t) + \dot\ell(\vartheta, t)]^2 \, y_t(\vartheta)^2}{b(\vartheta, t)^2} \, \mathrm{d}t, \qquad \vartheta \in \Theta.$$

Define as well the function

$$G(\vartheta, \vartheta_0) = \int_0^T \frac{[a(\vartheta, t) - a(\vartheta_0, t) + \ell(\vartheta, t) - \ell(\vartheta_0, t)]^2 y_t(\vartheta_0)^2}{b(\vartheta, t)^2} \mathrm{d}t.$$

Note that the Fisher information does not depend on $\sigma(\cdot)$.

Conditions $\hat{\mathscr{L}}$.

$\mathscr{L}_1$. *The functions* $f(\vartheta,t), a(\vartheta,t), b(\vartheta,t), t \in [0,T], \vartheta \in \Theta$ *and* $\sigma(t), t \in [0,T]$ *have continuous derivatives on* t.

$\mathscr{L}_2$. *The functions* $f(\vartheta,t), b(\vartheta,t), t \in [0,T], \vartheta \in \Theta$ *and* $\sigma(t), 0 \leq t \leq T$ *are separated from zero.*

$\mathscr{L}_3$. *The functions* $f(\vartheta,t), a(\vartheta,t), t \in [0,T], \vartheta \in \Theta$ *are two times continuously differentiable on* ϑ.

$\mathscr{L}_5$. *The Fisher information is positive*

$$\inf_{\vartheta \in \Theta} \mathrm{I}(\vartheta) > 0.$$

$\mathscr{L}_6$. *For any* $\nu > 0$

$$g(\nu) \equiv \inf_{\vartheta_0 \in \Theta} \inf_{|\vartheta - \vartheta_0| > \nu} G(\vartheta, \vartheta_0) > 0.$$

The mean squared estimation error satisfies the Hajek's minimax lower bound: for any $\vartheta_0 \in \Theta$ and any estimator $\bar\vartheta_\varepsilon$

$$\lim_{\nu\to 0} \varliminf_{\varepsilon\to 0} \sup_{|\vartheta-\vartheta_0|\leq\nu} \psi_\varepsilon^{-2} \mathbf{E}_\vartheta \left|\bar\vartheta_\varepsilon - \vartheta\right|^2 \geq \mathrm{I}\left(\vartheta_0\right)^{-1}.$$

The estimator ϑ_ε^* is called asymptotically efficient if for any $\vartheta_0 \in \Theta$

$$\lim_{\nu\to 0} \lim_{\varepsilon\to 0} \sup_{|\vartheta-\vartheta_0|\leq\nu} \psi_\varepsilon^{-2} \mathbf{E}_\vartheta \left|\vartheta_\varepsilon^* - \vartheta\right|^2 = \mathrm{I}\left(\vartheta_0\right)^{-1}.$$

Theorem

Assume that conditions $\hat{\mathscr{L}}$ are satisfied. Then the MLE $\hat\vartheta_\varepsilon$ and BE $\tilde\vartheta_\varepsilon$ are uniformly consistent, asymptotically normal

$$\frac{(\hat\vartheta_\varepsilon - \vartheta_0)}{\psi_\varepsilon} \Longrightarrow \zeta(\vartheta_0) \sim \mathcal{N}\left(0, \mathrm{I}(\vartheta_0)^{-1}\right), \ \frac{(\tilde\vartheta_\varepsilon - \vartheta_0)}{\psi_\varepsilon} \Longrightarrow \zeta(\vartheta_0),$$

the convergence of moments holds,

$$\mathbf{E}_{\vartheta_0}\left|\frac{\hat\vartheta_\varepsilon - \vartheta_0}{\psi_\varepsilon}\right|^p \to \mathbf{E}_{\vartheta_0}|\zeta(\vartheta_0)|^p, \quad \mathbf{E}_{\vartheta_0}\left|\frac{\tilde\vartheta_\varepsilon - \vartheta_0}{\psi_\varepsilon}\right|^p \to \mathbf{E}_{\vartheta_0}|\zeta(\vartheta_0)|^p,$$

for any $p > 0$, and both estimators are asymptotically efficient.

Let us denote $\gamma_0\left(\vartheta, t\right) = b\left(\vartheta, t\right)\sigma(t)/f\left(\vartheta, t\right)$ and $\vartheta_u = \vartheta_0 + \psi_\varepsilon\, u$

Lemma

Let the condition $\varepsilon{\psi_\varepsilon}^{-1} \rightarrow 0$ *hold . Then, for any* $t_0 \in (0, T]$*,*

$$\sup_{t_0 \leq t \leq T} \left| \frac{\gamma\left(\vartheta, t\right)}{\varepsilon\psi_\varepsilon} - \gamma_0\left(\vartheta, t\right) \right| = O\left(\frac{\varepsilon}{\psi_\varepsilon}\right).$$

The function $\gamma\left(\vartheta, t\right)$ in K-B filter is replaced by $\varepsilon\psi_\varepsilon\gamma_0\left(\vartheta, t\right)$.

Recall that $M(\vartheta,t) = f(\vartheta,t)\, m(\vartheta,t)$ and If $1/3 < \delta < 1$, then $\varphi_\varepsilon = \psi_\varepsilon$

$$\delta_M(t) = \frac{\varepsilon}{\psi_\varepsilon}\ \frac{a(\vartheta,t) - a(\vartheta_0,t) + \ell(\vartheta,t) - \ell(\vartheta_0,t)}{B(\vartheta,t)} f(\vartheta_0,t)\, y_t(\vartheta_0)$$

Lemma

For any $u \in \mathbb{U}_\varepsilon$ and any $t \in (0,T]$ the relation

$$\frac{M(\vartheta_u,t) - M(\vartheta_0,t)}{\sigma(t)} = \varepsilon\, u \frac{[\dot a(\vartheta_0,t) + \dot\ell(\vartheta_0,t)] y_t(\vartheta_0)}{b(\vartheta_0,t)} (1 + o(1)),$$

holds.

One-step MLE-process.

The learning interval is $[0, \tau_\varepsilon]$, where $\tau_\varepsilon = \varepsilon/\psi_\varepsilon$ and the preliminary MDE is

$$\check{\vartheta}_{\tau_\varepsilon} = \arg\inf_{\vartheta\in\Theta} \int_0^{\tau_\varepsilon} \left[X_t - x_t\left(\vartheta\right)\right]^2 \mathrm{d}t.$$

Note that $\sqrt{\tau_\varepsilon}\psi_\varepsilon^{-1} = (\varepsilon/\psi_\varepsilon^3)^{1/2} \to \infty$. Therefore

$$\varepsilon^{\frac{\delta-1}{2}}(\check{\vartheta}_{\tau_\varepsilon} - \vartheta_0) \Longrightarrow \mathcal{N}\left(0, D\left(\vartheta_0\right)\right).$$

Introduce the random process $M(\check{\vartheta}_{\tau_\varepsilon}, t), \tau_\varepsilon \leq t \leq T$ by

$$\mathrm{d}M(\check{\vartheta}_{\tau_\varepsilon}, t) = A^*(\check{\vartheta}_{\tau_\varepsilon}, t)M(\check{\vartheta}_{\tau_\varepsilon}, t)\mathrm{d}t + \frac{\psi_\varepsilon}{\varepsilon}\tilde{B}(\check{\vartheta}_{\tau_\varepsilon}, t)\,\mathrm{d}X_t.$$

subject to initial value $M(\check{\vartheta}_{\tau_\varepsilon}, \tau_\varepsilon)$.

Here

$$A^*(\check\vartheta_{\tau_\varepsilon}, t) = a^*(\check\vartheta_{\tau_\varepsilon}, t) - \frac{\psi_\varepsilon}{\varepsilon}\tilde B^*(\check\vartheta_{\tau_\varepsilon}, t),$$

$$\tilde B(\check\vartheta_{\tau_\varepsilon}, t) = \frac{\gamma_*(\check\vartheta_{\tau_\varepsilon}, t) f(\check\vartheta_{\tau_\varepsilon}, t)^2}{\sigma(t)^2}.$$

Define the Fisher information function

$$\mathrm{I}_{\tau_\varepsilon}^t(\vartheta) = \int_{\tau_\varepsilon}^t \frac{\dot a^*(\vartheta, s)^2 y_s(\vartheta)^2}{b(\vartheta, s)^2}\mathrm{d}s, \qquad \tau_\varepsilon \le t \le T.$$

Introduce the One-step MLE-process $\vartheta^\star_{t,\varepsilon}, \tau_\varepsilon < t \le T$.

$$\vartheta^\star_{t,\varepsilon} = \check\vartheta_{\tau_\varepsilon} + \frac{\psi_\varepsilon}{\mathrm{I}_{\tau_\varepsilon}^t(\check\vartheta_{\tau_\varepsilon})}\int_{\tau_\varepsilon}^t \frac{\dot a^*(\check\vartheta_{\tau_\varepsilon}, s) y_s(\check\vartheta_{\tau_\varepsilon})}{\varepsilon\ b(\check\vartheta_{\tau_\varepsilon}, s)\sigma(s)}\left[\mathrm{d}X_s - M(\check\vartheta_{\tau_\varepsilon}, s)\mathrm{d}s\right].$$

Theorem

Let the conditions $\mathscr{L}$ hold and $\delta \in (1/3, 1/2)$. Then the estimator process $\vartheta^\star_{t,\varepsilon}, \tau_\varepsilon \leq t \leq T$ is uniformly consistent, asymptotically normal

$$\psi_\varepsilon^{-1}(\vartheta^\star_{t,\varepsilon} - \vartheta_0) \Longrightarrow \mathcal{N}(0, \mathrm{I}_0^t(\vartheta_0)^{-1})$$

If the learning interval is fixed, then this convergence is valid for $\delta \in (1/3, 1)$.

Adaptive filter.

Suppose that the learning interval is $[0,\tau]$, $0<\tau<T$ and we have the MDE $\check\vartheta_\varepsilon$ constructed by the observations $X^\tau=(X_t, 0\leq t\leq \tau)$. This allow us to admit the values $\delta\in(1/3,1)$ for the construction of the One-step MLE-process

$$\vartheta^\star_{t,\varepsilon}=\check\vartheta_\varepsilon+\mathrm{I}_\tau^t(\check\vartheta_\varepsilon)^{-1}\psi_\varepsilon\int_\tau^t\frac{\dot a^*(\check\vartheta_\varepsilon,s)y_s(\check\vartheta_\varepsilon)}{\varepsilon\, b(\check\vartheta_\varepsilon,s)\sigma(s)}\left[\mathrm{d}X_s-M(\check\vartheta_\varepsilon,s)\mathrm{d}s\right].$$

The adaptive filter is defined by the equation

$$\mathrm{d}m^\star_{t,\varepsilon}=\left[a(\vartheta^\star_{t,\varepsilon},t)-\frac{\psi_\varepsilon}{\varepsilon}\frac{b(\vartheta^\star_{t,\varepsilon},t)f(\vartheta^\star_{t,\varepsilon},t)}{\sigma(t)}\right]m^\star_{t,\varepsilon}\mathrm{d}t+\frac{\psi_\varepsilon}{\varepsilon}\frac{b(\vartheta^\star_{t,\varepsilon},t)}{\sigma(t)}\mathrm{d}X_t$$

for the values $\tau_*\leq t\leq T$, where $\tau_*\in(\tau,T)$. It is asymptotically efficient.

1 Adaptive filtering

2 Different levels of noise

3 The same level of noise ($\psi_\varepsilon = \varepsilon$)

4 Low noise in observations ($\psi_\varepsilon = 1$)

5 Different Levels ($\psi_\varepsilon = \varepsilon^\delta, 1/3 < \delta < 1$)

6 Different Levels ($\delta = 1/3$)

7 Different Levels ($0 < \delta < 1/3$)

The system is

$$\mathrm{d}X_t = f(\vartheta, t)\, Y_t \mathrm{d}t + \varepsilon\sigma(t)\, \mathrm{d}W_t, \qquad X_0 = 0, \qquad 0 \le t \le T,$$
$$\mathrm{d}Y_t = a(\vartheta, t)\, Y_t \mathrm{d}t + \varepsilon^{1/3} b(\vartheta, t)\, \mathrm{d}V_t, \quad Y_0 = y_0,$$

We have the representation $(\vartheta_u = \vartheta_0 + \varphi_\varepsilon u)$

$$\frac{\delta_M(t)}{\varepsilon} = \frac{a(\vartheta_u, t) - a(\vartheta_0, t) + \ell(\vartheta_u, t) - \ell(\vartheta_0, t)}{\varepsilon^{1/3} B(\vartheta, t)} f(\vartheta_0, t)\, y_t(\vartheta_0)$$
$$+ \frac{[f(\vartheta_u, t)\, b(\vartheta_u, t) - f(\vartheta_0, t)\, b(\vartheta_0, t)]}{\varepsilon^{1/3}\sqrt{2B(\vartheta, t)}}\, \xi_{t,\varepsilon}.$$

and therefore $\varphi_\varepsilon = \varepsilon^{1/3}$. Hence

$$\frac{\delta_M(t)}{\varepsilon} = u\, \frac{\dot a(\vartheta_0, t) + \dot\ell(\vartheta_0, t)}{B(\vartheta_0, t)} f(\vartheta_0, t)\, y_t(\vartheta_0)$$
$$+ u\, \frac{\left[\dot f(\vartheta_0, t)\, b(\vartheta_0, t) + f(\vartheta_0, t)\, \dot b(\vartheta_0, t)\right]}{\sqrt{2B(\vartheta_0, t)}}\, \xi_{t,\varepsilon}.$$

The same MDE with the learning interval $[0, \tau_\varepsilon]$, $\tau_\varepsilon = \varepsilon^{2/3}$ is consistent and asymptotically normal

$$\varepsilon^{-1/3}(\check{\vartheta}_{\tau_\varepsilon} - \vartheta_0) \Longrightarrow \mathcal{N}\left(0, D\left(\vartheta_0\right)\right).$$

Fisher information

$$\begin{aligned}\mathrm{I}\left(\vartheta\right) = \int_0^T \frac{[\dot{f}\left(\vartheta,s\right) b\left(\vartheta,s\right) + f\left(\vartheta,s\right)\dot{b}\left(\vartheta,s\right)]^2}{2f\left(\vartheta,s\right) b\left(\vartheta,s\right)\sigma\left(s\right)}\,\mathrm{d}s \\ + \int_\tau^t \frac{[\dot{a}\left(\vartheta,s\right) + \dot{\ell}\left(\vartheta,s\right)]^2 y_s\left(\vartheta\right)^2}{b\left(\vartheta,s\right)^2}\,\mathrm{d}s.\end{aligned}$$

The MLE, BE are consistent and asymptotically normal

$$\varepsilon^{-1/3}\left(\hat{\vartheta}_\varepsilon - \vartheta_0\right) \Longrightarrow \hat{\eta}_T \sim \mathcal{N}\left(0, \mathrm{I}\left(\vartheta_0\right)^{-1}\right), \quad \varepsilon^{-1/3}\left(\tilde{\vartheta}_\varepsilon - \vartheta_0\right) \Longrightarrow \hat{\eta}_T$$

The One-step MLE-process is

$$\vartheta^{\star}_{t,\varepsilon} = \check{\vartheta}_{\tau_\varepsilon} + \frac{\mathrm{I}^t_{\tau_\varepsilon}\left(\check{\vartheta}_{\tau_\varepsilon}\right)^{-1}}{\varepsilon^{4/3}} \int_{\tau_\varepsilon}^{t} \frac{\dot{M}(\check{\vartheta}_{\tau_\varepsilon}, s)}{\sigma\left(s\right)^2} \left[\mathrm{d}X_s - M(\check{\vartheta}_{\tau_\varepsilon}, s)\mathrm{d}s\right]$$

The estimator is asymptotically normal

$$\varepsilon^{-1/3}\left(\vartheta^{\star}_{t,\varepsilon} - \vartheta_0\right) \Longrightarrow \eta^{\star}_t\left(\vartheta_0\right) \sim \mathcal{N}\left(0, \mathrm{I}^t_0\left(\vartheta_0\right)^{-1}\right).$$

The limit random variable $\eta_t^\star(\vartheta_0)$ can be written with the help of two independent Wiener processes

$$\eta_t^\star(\vartheta_0) = \mathrm{I}_0^t(\vartheta_0)^{-1}\left\{\int_0^t \frac{[\dot f(\vartheta_0,s)\,b(\vartheta_0,s) + f(\vartheta_0,s)\,\dot b(\vartheta_0,s)]}{\sqrt{2f(\vartheta_0,s)\,b(\vartheta_0,s)\,\sigma(s)}}\,\mathrm{d}w_1(s)\right.$$
$$\left. + \int_0^t \frac{[\dot a(\vartheta_0,s) + \dot\ell(\vartheta_0,s)]y_s(\vartheta_0)}{b(\vartheta_0,s)}\,\mathrm{d}w_2(s)\right\}.$$

Adaptive filter

The adaptive filter is

$$\mathrm{d}m^{\star}_{t,\varepsilon} = \left[a^{*}\left(\vartheta^{\star}_{t,\varepsilon}, t\right) - \varepsilon^{-2/3} B\left(\vartheta^{\star}_{t,\varepsilon}, t\right)\right] m^{\star}_{t,\varepsilon}\mathrm{d}t + \varepsilon^{-2/3} B\left(\vartheta^{\star}_{t,\varepsilon}, t\right)\mathrm{d}X_t,$$

where $B(\vartheta, t) = f(\vartheta, t)\, b(\vartheta, t)\, \sigma(t)^{-1}$ and $\tau_\varepsilon < t \leq T$.

The error of approximation is

$$\frac{m^{\star}_{t,\varepsilon} - m(\vartheta_0, t)}{\varepsilon^{1/3}}$$

$$= \left[\frac{\dot{b}(\vartheta_0, t)\sqrt{\sigma(t)}}{\sqrt{2b(\vartheta_0, t)\, f(\vartheta_0, t)}}\,\xi_{t,\varepsilon} - \frac{\dot{f}(\vartheta_0, t)\;\, y_t(\vartheta_0)}{f(\vartheta_0, t)}\right]\eta^{\star}_{t,\varepsilon}\left(1 + o(1)\right)$$

$$\Longrightarrow \left[\frac{\dot{b}(\vartheta_0, t)\sqrt{\sigma(t)}}{\sqrt{2b(\vartheta_0, t)\, f(\vartheta_0, t)}}\,\xi_{t} - \frac{\dot{f}(\vartheta_0, t)\;\, y_t(\vartheta_0)}{f(\vartheta_0, t)}\right]\eta^{\star}_{t}(\vartheta_0).$$

1. Adaptive filtering
2. Different levels of noise
3. The same level of noise ($\psi_\varepsilon = \varepsilon$)
4. Low noise in observations ($\psi_\varepsilon = 1$)
5. Different Levels ($\psi_\varepsilon = \varepsilon^\delta, 1/3 < \delta < 1$)
6. Different Levels ($\delta = 1/3$)
7. Different Levels ($0 < \delta < 1/3$)

Recall the partially observed system

$$\begin{aligned} \mathrm{d}X_t &= f(\vartheta, t)\, Y_t \mathrm{d}t + \varepsilon\sigma(t)\,\mathrm{d}W_t, \qquad X_0 = 0, \qquad 0 \le t \le T, \\ \mathrm{d}Y_t &= a(\vartheta, t)\, Y_t \mathrm{d}t + \varepsilon^{\delta} b(\vartheta, t)\,\mathrm{d}V_t, \quad Y_0 = y_0, \end{aligned}$$

where $0 < \delta < 1/3$. The representation is

$$\begin{aligned} &\frac{M(\vartheta_u, t) - M(\vartheta_0, t)}{\varepsilon} \\ &\qquad = \sqrt{\frac{\psi_\varepsilon}{\varepsilon}}\ \frac{\left[f(\vartheta_u, t)\, b(\vartheta_u, t) - f(\vartheta_0, t)\, b(\vartheta_0, t)\right]}{\sqrt{2B(\vartheta, t)}}\xi_{t,\varepsilon} \\ &\qquad = u\ \varphi_\varepsilon \sqrt{\frac{\psi_\varepsilon}{\varepsilon}}\ \frac{\left[\dot f(\vartheta_0, t)\, b(\vartheta_0, t) + f(\vartheta_0, t)\, \dot b(\vartheta_0, t)\right]}{\sqrt{2B(\vartheta, t)}}\xi_{t,\varepsilon}. \end{aligned}$$

Therefore $\varphi_\varepsilon = \sqrt{\frac{\varepsilon}{\psi_\varepsilon}} = \varepsilon^{\frac{(1-\delta)}{2}}$.

MDE The MDE is the same

$$\check{\vartheta}_{\tau,\varepsilon} = \arg\inf_{\vartheta\in\Theta} \int_0^\tau \left[X_t - x_t\left(\vartheta\right)\right]^2 \mathrm{d}t.$$

If the interval $[0, \tau]$ is fixed, then

$$\varepsilon^{-\delta}(\check{\vartheta}_{\tau,\varepsilon} - \vartheta_0) \Longrightarrow \mathcal{N}\left(0, D_1\left(\vartheta_0\right)\right).$$

If the interval $[0, \tau_\varepsilon]$ is vanishing, $\tau_\varepsilon = \varepsilon^{1-\delta}$, then

$$\varepsilon^{\frac{(\delta-1)}{2}}(\check{\vartheta}_{\tau,\varepsilon} - \vartheta_0) \Longrightarrow \mathcal{N}\left(0, D_2\left(\vartheta_0\right)\right).$$

MLE and BE Under regularity conditions

$$\sqrt{\frac{\psi_\varepsilon}{\varepsilon}}(\hat\vartheta_\varepsilon - \vartheta_0) \Longrightarrow \zeta\left(\vartheta_0\right) \sim \mathcal{N}\left(0, \mathrm{I}\left(\vartheta_0\right)^{-1}\right),$$
$$\sqrt{\frac{\psi_\varepsilon}{\varepsilon}}\left(\tilde\vartheta_\varepsilon - \vartheta_0\right) \Longrightarrow \zeta\left(\vartheta_0\right)$$

and polynomial moments converge too.

Here the Fisher information is

$$\mathrm{I}\left(\vartheta\right) = \int_0^T \frac{\left[\dot f\left(\vartheta,t\right) b\left(\vartheta,t\right) + f\left(\vartheta,t\right)\dot b\left(\vartheta,t\right)\right]^2}{2b\left(\vartheta,t\right) f\left(\vartheta,t\right)\sigma\left(t\right)}\mathrm{d}t.$$

One-step MLE-process The One-step MLE-process $\vartheta^\star_{t,\varepsilon}, \tau_\varepsilon < t \leq T$ is

$$\vartheta^\star_{t,\varepsilon} = \check\vartheta_{\tau_\varepsilon} + \frac{\mathrm{I}^t_{\tau_\varepsilon}\left(\check\vartheta_{\tau_\varepsilon}\right)^{-1}}{\varepsilon\,\psi_\varepsilon}\int_{\tau_\varepsilon}^{t}\frac{\dot M(\check\vartheta_{\tau_\varepsilon},s)}{\sigma\left(s\right)^2}\left[\mathrm{d}X_s - M(\check\vartheta_{\tau_\varepsilon},s)\mathrm{d}s\right].$$

The random processes $M(\check\vartheta_{\tau_\varepsilon},s)$ and $\dot M(\check\vartheta_{\tau_\varepsilon},s)$ are defined by the same equations as before and $\tau_\varepsilon = \varepsilon^{1-\delta}$.

We have

$$\frac{\vartheta^\star_{t,\varepsilon} - \vartheta_0}{\varphi_\varepsilon} \Longrightarrow \eta^\star_t\left(\vartheta_0\right) = \mathrm{I}^t_0\left(\vartheta_0\right)^{-1}\int_0^t \frac{\dot S\left(\vartheta_0,s\right)}{\sqrt{2S\left(\vartheta_0,s\right)\sigma\left(s\right)}}\,\mathrm{d}w\left(s\right)$$

Here $S\left(\vartheta,s\right) = f\left(\vartheta,s\right)b\left(\vartheta,s\right)$.

Adaptive filter.

Case $\delta \in (1/5, 1/3)$. The adaptive filter $m^{\star}_{t,\varepsilon}, \tau < t \leq T$ is

$$\mathrm{d}m^{\star}_{t,\varepsilon} = \left[a\left(\vartheta^{\star}_{t,\varepsilon}, t\right) - \varepsilon^{-1+\delta} B\left(\vartheta^{\star}_{t,\varepsilon}, t\right)\right] m^{\star}_{t,\varepsilon}\mathrm{d}t + \varepsilon^{-1+\delta}\frac{b\left(\vartheta^{\star}_{t,\varepsilon}, t\right)}{\sigma\left(t\right)}\mathrm{d}X_t,$$

and the error of estimation has the following asymptotics

$$\frac{m^{\star}_{t,\varepsilon} - m\left(\vartheta_0, t\right)}{\varepsilon^{(5\delta-1)/2}} \Longrightarrow \frac{\dot{b}\left(\vartheta_0, t\right)\sqrt{\sigma\left(t\right)}}{\sqrt{2b\left(\vartheta_0, t\right) f\left(\vartheta_0, t\right)}}\ \eta^{\star}_t\left(\vartheta_0\right)\ \xi_t,$$

where $B\left(\vartheta, t\right) = b\left(\vartheta, t\right) f\left(\vartheta, t\right) \sigma\left(t\right)^{-1}$ and $\xi_t \sim \mathcal{N}\left(0, 1\right)$ and

$$\eta^{\star}_t\left(\vartheta_0\right) = \mathrm{I}^t_{\tau}\left(\vartheta_0\right)^{-1}\int_{\tau}^{t}\frac{\dot{f}\left(\vartheta_0, s\right) b\left(\vartheta_0, s\right) + f\left(\vartheta_0, s\right)\dot{b}\left(\vartheta_0, s\right)}{\sqrt{2f\left(\vartheta_0, s\right) b\left(\vartheta_0, s\right)\sigma\left(s\right)}}\mathrm{d}w\left(s\right).$$

K. (2025) *Hidden Markov Processes and Adaptive Filtering*. Springer, Cham

Thank you for your attention!