HMM in which Higher Noise and Fewer of Observations Improve Parameter Estimators (Revisited)

Yury A. Kutoyants

Le Mans University, France

Workshop on Statistics of Stochastic Processes

Le Mans — 18 September, 2026

1 Hidden Markov Processes

2 MDE.

3 MLE and BE.

4 One-step MLE

5 Adaptive filters.

Kalman-Bucy filter.

We are given a couple of equations

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

where $W_t, 0 \le t \le T$ and $V_t, 0 \le t \le T$ are independent Wiener processes and the observations are $X^T = (X_t, 0 \le t \le T)$. The process $Y^T = (Y_t, 0 \le t \le T)$ is hidden. The functions $f(\vartheta, t), \sigma(t), a(\vartheta, t), b(\vartheta, t), t \in [0, T]$ are supposed to be known, smooth and the parameter $\vartheta \in \Theta \subset \mathcal{R}^d$ is unknown. The function $\sigma(t)^2 \ge c > 0$. The initial value y_0 is deterministic and known.

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

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

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

with initial values $m_0 = \mathbf{E}_\vartheta\left(Y_0 | X_0\right)$ and $\gamma_0 = \mathbf{E}_\vartheta\left(Y_0 - m(\vartheta, 0)\right)^2$.

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

Solution: *Estimate* ϑ *and substitute this estimator in the K-B equations.*

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 $\left\{m\left(\vartheta, s\right),\ 0 \le s \le t\right\},\ \vartheta \in \Theta$ and $\left\{\gamma\left(\vartheta, s\right),\ 0 \le s \le t\right\},\ \vartheta \in \Theta$ 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 and BE is 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* $\vartheta_t^\star, \tau < t \leq T$
3. *As approximation of* $m(\vartheta_0, t)$ *we propose* $m_t^\star$ *obtained with the help of K-B equations, where* ϑ *is replaced by* $\vartheta_t^\star, \tau < t \leq T$
4. *Estimate the error* $\mathbf{E}_{\vartheta_0} [m_t^\star - m(\vartheta_0, t)]^2$ *and show its optimality.*

Basic Model. Different Levels.

Suppose that

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

Observations are $X^T = (X_t, 0 \le t \le T)$ and the asymptotic $\varepsilon \to 0, \psi_\varepsilon \to 0$. The properties of MLEs are defined by the asymptotic of the integral $(M(\vartheta, t) = f(\vartheta, t)\, m(\vartheta, t))$

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

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

$$\frac{M(\vartheta,t)-M(\vartheta_0,t)}{\varepsilon\sigma(t)}$$

$$=\frac{y_0\left[f(\vartheta,0)-f(\vartheta_0,0)\right]}{\varepsilon\,\sigma(t)}\exp\left(\int_0^t\left[a(\vartheta,s)-\frac{f(\vartheta,s)^2\gamma(\vartheta,s)}{\varepsilon^2\sigma(s)^2}\right]\mathrm{d}s\right)$$

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

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

where $\xi_{t,\varepsilon}\Longrightarrow\xi_t\;\sim\;\mathcal{N}(0,1)$. The random variables $\xi_t, 0<t\leq T$ are independent.

Two cases. **A:** $\dot{f}(\vartheta, 0) = 0$ for all $\vartheta \in \Theta$ (we obtain results from my first talk) and **B:** $\dot{f}(\vartheta, 0) \neq 0$ for all $\vartheta \in \Theta$.

1. **A.** $\psi_\varepsilon = 1, (\delta = 0),$ $\varphi_\varepsilon = \sqrt{\varepsilon},$ $\mathrm{I}_1(\vartheta_0),$
2. **A.** $\psi_\varepsilon = \varepsilon^\delta, (0 < \delta < 1/3\ ),$ $\varphi_\varepsilon = \varepsilon^{\frac{1-\delta}{2}},$ $\mathrm{I}_1(\vartheta_0)\ ,$
3. **A.** $\psi_\varepsilon = \varepsilon^{1/3}, (\delta = 1/3),$ $\varphi_\varepsilon = \varepsilon^{\frac{1}{3}},$ $\mathrm{I}_3(\vartheta_0) = \mathrm{I}_1(\vartheta_0) + \mathrm{I}_2(\vartheta_0),$

   .
4. **A.** $\psi_\varepsilon = \varepsilon^\delta, (1/3 < \delta < 1),$ $\varphi_\varepsilon = \varepsilon^\delta,$ $\mathrm{I}_2(\vartheta_0),$
5. **A.** $\psi_\varepsilon = \varepsilon, (\delta = 1),$ $\varphi_\varepsilon = \varepsilon,$ $\mathrm{I}_4(\vartheta_0).$
6. **B.** $\psi_\varepsilon = \varepsilon^\delta, (0 < \delta < 1\ ),$ $\varphi_\varepsilon = \varepsilon^{\frac{1+\delta}{2}},$ $\mathrm{I}_0(\vartheta_0).$

In all cases

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

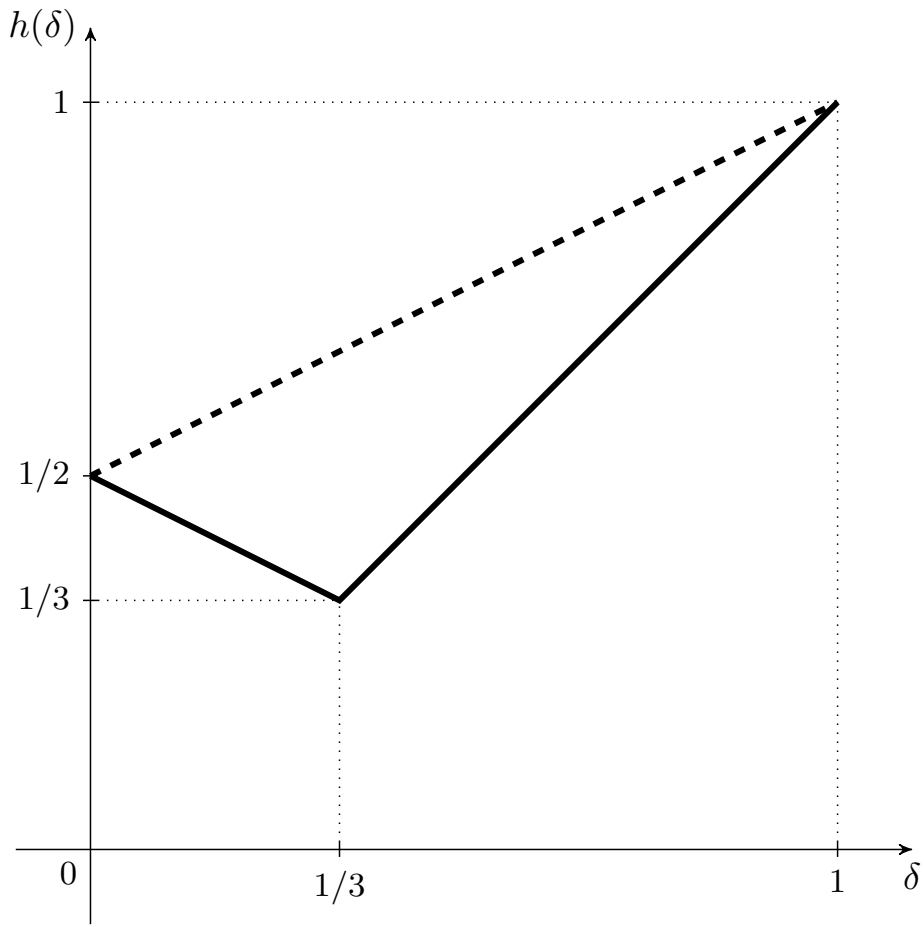


Figure: Function $h(\delta)$, **A** (continuous line) and **B** (dashed line)

MLE: $\delta \in \left(0, \frac{1}{3}\right)$ **the higher noise $\rightarrow$ the smaller error of estimation!**

Case C.

The parameter $\vartheta = (\theta_0, \theta_1, \theta_2)$, and the functions are $f(\theta_0, t)$, $b(\theta_1, t)$, $a(\theta_2, t)$.

The MLE $\hat{\vartheta}_\varepsilon$ is asymptotically normal

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

Here the normalizing matrix is

$$\varphi_\varepsilon = \begin{pmatrix} \left(\varepsilon\psi_\varepsilon\right)^{1/2} & 0 & 0 \\ 0, & \left(\varepsilon/\psi_\varepsilon\right)^{1/2} & 0 \\ 0, & 0, & \psi_\varepsilon^{1/2} \end{pmatrix}.$$

The Fisher information matrix $\mathbf{I}\left(\vartheta_0\right)_{3\times 3}$ is diagonal

$$\mathrm{I}_0\left(\vartheta\right) = \frac{y_0^2 \dot{f}\left(\theta_0, 0\right)^2}{f\left(\theta_0, 0\right) b\left(\theta_1, 0\right) \sigma\left(0\right)}, \qquad \mathrm{I}_1\left(\vartheta\right) = \int_0^T \frac{\dot{b}\left(\theta_1, t\right)^2 f\left(\theta_0, t\right)}{2 b\left(\theta_1, t\right) \sigma\left(t\right)} \, \mathrm{d}t,$$

$$\mathrm{I}_2\left(\vartheta\right) = \int_0^T \frac{\dot{a}\left(\theta_2, t\right)^2 \, y_t\left(\vartheta\right)^2}{b\left(\theta_1, t\right)^2} \, \mathrm{d}t.$$

1 Hidden Markov Processes

2 **MDE.**

3 **MLE and BE.**

4 One-step MLE

5 Adaptive filters.

MDE

. The observations X^T are considered as a deterministic system

$$\frac{\mathrm{d}x_t(\vartheta)}{\mathrm{d}t} = f(\vartheta, t)\, y_t(\vartheta), \qquad x_0(\vartheta) = 0,$$
$$\frac{\mathrm{d}y_t(\vartheta)}{\mathrm{d}t} = a(\vartheta, t)\, y_t(\vartheta), \qquad y_0(\vartheta) = y_0, \quad 0 \le t \le T.$$

perturbed by two small Gaussian noises. The minimum distance estimator (MDE) $\check{\vartheta}_{\tau,\varepsilon}$ constructed by the observations X^τ is defined by the equation

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

Remark that

$$x_t(\vartheta) = y_0 \int_0^t f(\vartheta, s) \exp\left(\int_0^s a(\vartheta, r)\,\mathrm{d}r\right) \mathrm{d}s, \qquad \vartheta \in \Theta.$$

We study this estimator in cases, **A** and **B**, and in two settings:

Remark.

If one says that in the case when the limit system does not depend on ϑ, then the consistent estimation is impossible, then this is wrong! For example, if

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

then the limit system

$$\frac{\mathrm{d}x_t}{\mathrm{d}t} = f(t)\, y_t, \quad x_0 = 0, \qquad \frac{\mathrm{d}y_t}{\mathrm{d}t} = a(t)\, y_t, \quad y_0,\ 0 \le t \le T$$

does not depend on ϑ, but the consistent and asymptotically efficient estimation of ϑ is nevertheless possible (under regularity conditions). This is a particular case of the Case **C**.

K. (2024) Volatility estimation of hidden Markov process and adaptive filtration. Stoch. Proc and Appl., 173, 104381.

Cases A and B. Fixed interval. Introduce the $d \times d$ matrix

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

Gaussian vector

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

where

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

and the function

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

Conditions $\mathscr{B}$.

$\mathscr{B}_1$. *The functions* $f(\vartheta, t), a(\vartheta, t), b(\vartheta, t) \in \mathcal{C}^2_\vartheta$ *for all* $\vartheta \in \Theta, t \in [0, \tau]$.

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

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

$\mathscr{B}_3$. *For any* $\nu > 0$ *we have* $\inf_{\vartheta_0 \in \Theta} \check{g}(\vartheta_0, \nu) > 0$.

Theorem

Let the conditions $\mathscr{B}$ hold. Then the MDE is consistent, asymptotically normal $(\phi_\varepsilon = \psi_\varepsilon = \varepsilon^\delta)$

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

and the polynomial moments converge.

Case A. Vanishing interval. $\delta \in (1/3, 1)$**.** We study

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

Denote $R\left(\vartheta\right) = y_0 \left[\dot{f}_t' \left(\vartheta, 0 \right) a \left(\vartheta, 0 \right) + f \left(\vartheta, 0 \right) \dot{a} \left(\vartheta, 0 \right) \right]$. Below $t_0 > 0$ is some (small) real from the interval $[0, T]$.

Conditions $\mathscr{C}$.

- $\mathscr{C}_0$. *The function* $\dot{f} \left(\vartheta, 0 \right) = 0$ *for all* $\vartheta \in \Theta$.
- $\mathscr{C}_1$. *The functions* $f \left(\vartheta, t \right), a \left(\vartheta, t \right) \in \mathcal{C}_\vartheta^2$ *for* $t \in [0, t_0]$.
- $\mathscr{C}_2$. *The functions* $f \left(\vartheta, t \right), \dot{f} \left(\vartheta, t \right), a \left(\vartheta, t \right), b \left(\vartheta, t \right), \sigma \left(t \right) \in \mathcal{C}_t^1$ *for* $t \in [0, t_0]$.
- $\mathscr{C}_3$. *The functions* $f \left(\vartheta, 0 \right), R \left(\vartheta \right)$ *and* $b \left(\vartheta, 0 \right)$ *are separated from zero uniformly on* $\vartheta \in \Theta$.

Let us denote

$$D^{\circ}(\vartheta) = \frac{130\, f(\vartheta,0)^2\, b(\vartheta,0)^2}{81\, R(\vartheta)^2}.$$

Below $\phi_\varepsilon = \varepsilon^{\delta} \ln \varepsilon^{-1}$ and $\tau_\varepsilon = [\ln \varepsilon^{-1}]^{-2}$.

Proposition

Let conditions $\mathscr{C}$ hold and $\delta \in (1/3, 1)$. Then the MDE $\check{\vartheta}_{\tau_\varepsilon}$ is consistent and asymptotically normal

$$\phi_\varepsilon^{-1}\left(\check{\vartheta}_{\tau_\varepsilon} - \vartheta_0\right) \Longrightarrow \zeta^{\circ}(\vartheta_0) \sim \mathcal{N}(0, D^{\circ}(\vartheta_0))$$

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

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

Case B. Vanishing interval. $\delta \in (1/3, 1)$.

Consider the same MDE with $\tau_\varepsilon \to 0$ but in the case **B**.

Conditions $\mathscr{D}$.

- $\mathscr{D}_0$. *The derivative* $|\dot{f}(\vartheta, 0)| \geq \kappa$.
- $\mathscr{D}_1$. *The functions* $f(\vartheta, t), a(\vartheta, t) \in \mathcal{C}^2_\vartheta$ *for* $t \in [0, t_0]$.
- $\mathscr{D}_2$. *The functions* $f(\vartheta, 0)$ *and* $b(\vartheta, 0)$ *are separated from 0.*
- $\mathscr{D}_3$. *The functions* $f(\vartheta, t), a(\vartheta, t), b(\vartheta, t), \sigma(t) \in \mathcal{C}^1_t$.

Let us denote

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

Proposition

Let conditions $\mathscr{D}$ hold, $\tau_\varepsilon = \varepsilon/\psi_\varepsilon$ and $\phi_\varepsilon = \varepsilon^{\frac{1+\delta}{2}}$. Then the MDE $\check{\vartheta}_{\tau_\varepsilon}$ is consistent, asymptotically normal

$$\phi_\varepsilon^{-1}\left(\check{\vartheta}_{\tau_\varepsilon} - \vartheta_0\right) \Longrightarrow \check{\zeta}\left(\vartheta_0\right) \sim \mathcal{N}\left(0, D^\circ\left(\vartheta_0\right)\right)$$

and the polynomial moments converge.

Remind that in the case of fixed interval the rate was $\phi_\varepsilon = \varepsilon^\delta$. Hence for the same model in the case **B** the rate of MDE constructed by all observations up to $[0, T]$ is worse

$$\frac{\varepsilon^\delta}{\varepsilon^{\frac{1+\delta}{2}}} = \varepsilon^{-\frac{1-\delta}{2}} \to \infty.$$

Conclusion: **less observations is better for estimation!**

1 Hidden Markov Processes

2 MDE.

3 **MLE and BE.**

4 One-step MLE

5 Adaptive filters.

MLE and BE

Once more we consider the cases **A** ($\dot{f}(\vartheta, 0) = 0$) and **B** ($\dot{f}(\vartheta, 0) \neq 0$). In the case **A** we study three situations: $\delta \in (1/3, 1)$, $\delta \in (0, 1/3)$,$\delta = 1/3$ separately because the LR has different limits with different normalizations.

Asymptotics of Riccati equation. Introduce the functions $\gamma_*(\vartheta, t) = \gamma(\vartheta, t) / (\varepsilon\psi_\varepsilon)$ and $\gamma_0(\vartheta, t) = b(\vartheta, t)\sigma(t) f(\vartheta, t)^{-1}$.

- For any $t_0 \in (0, T]$,

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

- For the values $[0, t_0]$, where t_0 is $o(1)$

$$\gamma_*(\vartheta, t) \approx \gamma^\circ(\vartheta, t) = \gamma_0(\vartheta, 0) \tanh\left(\frac{\psi_\varepsilon}{\varepsilon}\Gamma(\vartheta)\, t\right).$$

Here $\Gamma(\vartheta) = f(\vartheta, 0)\, b(\vartheta, 0)\sigma(0)^{-1}$.

Case **A**, $\psi_\varepsilon = \varepsilon^\delta$, $\delta \in (1/3, 1)$.

Then the Fisher information matrix is ($\ell(\vartheta, t) = [\ln f(\vartheta, t)]'_t$)

$$\mathbf{I}_2(\vartheta) = \int_0^T \frac{[\dot{a}(\vartheta, t) + \dot{\ell}(\vartheta, t)][\dot{a}(\vartheta, t) + \dot{\ell}(\vartheta, t)]^\top}{b(\vartheta, t)^2} y_t(\vartheta)^2 \, \mathrm{d}t, \ \vartheta \in \Theta.$$

Define as well the function

$$G_2(\vartheta, \vartheta_0) = \int_0^T \frac{[a(\vartheta, t) - a(\vartheta_0, t) + \ell(\vartheta, t) - \ell(\vartheta_0, t)]^2 y_t(\vartheta_0)^2}{b(\vartheta, t)^2} \mathrm{d}t.$$

Conditions $\mathscr{E}$.

$\mathscr{E}_0$. *The equality* $\dot{f}(\vartheta, 0) = 0$ *is valid for all* $\vartheta \in \Theta$.

$\hat{\mathscr{E}}_1$. *The functions* $f(\vartheta, t), a(\vartheta, t), b(\vartheta, t), \sigma(t) \in \mathcal{C}_t^1$.

$\mathscr{E}_2$. *The functions* $f(\vartheta, t), b(\vartheta, t)$ *are separated from zero.*

$\mathscr{E}_3$. *The functions* $f(\vartheta, t), a(\vartheta, t) \in \mathcal{C}_\vartheta^2$.

$\mathscr{E}_4$. *The Fisher information matrix is uniformly non degenerate*

$$\inf_{\vartheta \in \Theta} \inf_{\|\mathrm{e}\|=1, \mathrm{e} \in \mathcal{R}^d} \mathrm{e}^\top \mathbf{I}_2(\vartheta)\, \mathrm{e} > 0.$$

$\mathscr{E}_5$. *For any* $\nu > 0$

$$\inf_{\vartheta_0 \in \Theta} \inf_{\|\vartheta - \vartheta_0\| > \nu} G_2(\vartheta, \vartheta_0) > 0.$$

The family of measures $\left\{\mathbf{P}_{\vartheta}^{(\varepsilon)}, \vartheta \in \Theta\right\}$ induced by the observation process X^T on the space $(\mathcal{C}[0,T], \mathfrak{B}_T)$ of continuous on $[0,T]$ functions is locally asymptotically normal (LAN). Therefore the mean squared estimation error satisfies the Hajek's minimax lower bound: for any $\vartheta_0 \in \Theta$ and any estimator $\bar{\vartheta}_\varepsilon$

$$\lim_{\nu\to 0}\varliminf_{\varepsilon\to 0}\sup_{\|\vartheta-\vartheta_0\|\le\nu}\psi_\varepsilon^{-2}\mathbf{E}_\vartheta\left\|\bar{\vartheta}_\varepsilon-\vartheta\right\|^2\ge\mathbf{E}_{\vartheta_0}\left\|\zeta\left(\vartheta_0\right)\right\|^2.$$

Here $\zeta\left(\vartheta_0\right)\sim\mathcal{N}\left(0,\mathbf{I}_2\left(\vartheta_0\right)^{-1}\right)$. The estimator ϑ_ε^* is called asymptotically efficient if for any $\vartheta_0 \in \Theta$

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

Theorem

Let conditions $\mathscr{E}$ hold. Then the MLE $\hat{\vartheta}_\varepsilon$ and BE $\tilde{\vartheta}_\varepsilon$ are uniformly consistent, asymptotically normal

$$\psi_\varepsilon^{-1}\left(\hat{\vartheta}_\varepsilon - \vartheta_0\right) \Longrightarrow \zeta\left(\vartheta_0\right), \qquad \psi_\varepsilon^{-1}\left(\tilde{\vartheta}_\varepsilon - \vartheta_0\right) \Longrightarrow \zeta\left(\vartheta_0\right),$$

the convergence of moments holds,

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

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

Forbiden functions $f(\vartheta, t)$ and $a(\vartheta, t)$

The condition of identifiability requires that $G_2(\vartheta, \vartheta_0) \neq 0$ for any $\vartheta \neq \vartheta_0$ and any ϑ_0. This condition is not fulfilled if the functions $a(\vartheta, t)$ and $f(\vartheta, t)$ satisfy the equation

$$\frac{f'_t(\vartheta, t)}{f(\vartheta, t)} = -a(\vartheta, t) + h(t), \qquad 0 \leq t \leq T,$$

where $h(t)$ is some continuous function which does not depend on ϑ. This means that

$$f(\vartheta, t) = H(t) \exp\left(-\int_0^t a(\vartheta, s)\,\mathrm{d}s\right), \quad H(t) = \exp\left(\int_0^t h(s)\,\mathrm{d}s\right).$$

Therefore for such functions $f(\vartheta, \cdot)$ and $a(\vartheta, \cdot)$ we have $G_2(\vartheta, \vartheta_0) = 0$ for all $\vartheta \in \Theta$. If the functions $f(\vartheta, \cdot)$ and $a(\vartheta, \cdot)$ satisfy this relation then this case is in some sense close to the considered below case with $\delta \in (0, 1/3)$ and needs a special study.

Case **A**, $\psi_\varepsilon = \varepsilon^\delta$, $\delta \in (0, 1/3)$. We have

$$\frac{M(\vartheta_u, t) - M(\vartheta_0, t)}{\varepsilon\sigma(t)}$$

$$= \varphi_\varepsilon \sqrt{\frac{\psi_\varepsilon}{\varepsilon}} \frac{\left[u^\top \dot{f}(\vartheta_0, t) b(\vartheta_0, t) + f(\vartheta_0, t) u^\top \dot{b}(\vartheta_0, t)\right]}{\sqrt{2b(\vartheta_0, t) f(\vartheta_0, t) \sigma(t)}} \xi_{t,\varepsilon} (1 + o(1))$$

Therefore $\varphi_\varepsilon = \sqrt{\varepsilon/\psi_\varepsilon} = \varepsilon^{(1-\delta)/2}$. Here $\xi_{t,\varepsilon} \Rightarrow \xi_t \sim$ i.i.d. $\mathcal{N}(0, 1)$.

Let us denote

$$K(\vartheta, t) = \dot{f}(\vartheta, t) b(\vartheta, t) + f(\vartheta, t) \dot{b}(\vartheta, t)$$

and Fisher information matrix

$$\mathbf{I}_1(\vartheta) = \int_0^T \frac{K(\vartheta, t) K(\vartheta, t)^\top}{2f(\vartheta, t) b(\vartheta, t) \sigma(t)} \mathrm{d}t.$$

Define the function

$$G_1(\vartheta,\vartheta_0)=\int_0^T\frac{\left[f(\vartheta,t)\,b(\vartheta,t)-f(\vartheta_0,t)\,b(\vartheta_0,t)\right]^2}{2f(\vartheta,t)\,b(\vartheta,t)\;\sigma(t)}\mathrm{d}t$$

Conditions $\mathscr{F}$

$\mathscr{F}_1$. $f(\vartheta,t),b(\vartheta,t),\sigma(t)\in\mathcal{C}_t^1$

$\mathscr{F}_2$. $f(\vartheta,t),b(\vartheta,t),\sigma(t)$ *are bound. and separ. from zero.*

$\mathscr{F}_3$. $f(\vartheta,t),b(\vartheta,t)\in\mathcal{C}_\vartheta^2$.

$\mathscr{F}_4$. *The Fisher information is uniformly positive defined*

$$\inf_{\vartheta\in\Theta}\inf_{\|\mathrm{e}\|=1,\mathrm{e}\in\mathcal{R}^d}\mathrm{e}^\top\mathbf{I}_1(\vartheta)\,\mathrm{e}>0.$$

$\mathscr{F}_5$. *For any* $\nu>0$

$$\inf_{\vartheta_0\in\Theta}\inf_{\|\vartheta-\vartheta_0\|>\nu}G_1(\vartheta,\vartheta_0)>0.$$

Theorem

Let the conditions $\mathscr{F}$ hold and $\delta \in (0, 1/3)$. Then the MLE and BE are uniformly on compacts consistent, asymptotically normal

$$\sqrt{\frac{\psi_\varepsilon}{\varepsilon}}\left(\hat{\vartheta}_\varepsilon - \vartheta_0\right) \Longrightarrow \zeta^*\left(\vartheta_0\right) \sim \mathcal{N}\left(0, \mathbf{I}_1\left(\vartheta_0\right)^{-1}\right),$$

$$\sqrt{\frac{\psi_\varepsilon}{\varepsilon}}\left(\tilde{\vartheta}_\varepsilon - \vartheta_0\right) \Longrightarrow \zeta^*\left(\vartheta_0\right),$$

the moments converge and the both estimators are asymptotically efficient.

The rate of convergence $\varphi_\varepsilon = \varepsilon^{\frac{1-\delta}{2}}$ means that

the higher noise $\rightarrow$ the smaller errors.

Because all information about ϑ is in the product $\psi_\varepsilon f\left(\vartheta, t\right) b\left(\vartheta, t\right)$.

Case **A**, $\psi_\varepsilon = \varepsilon^{1/3}$, $\delta = 1/3$.

The Fisher information matrix is

$$\mathbf{I}_3(\vartheta_0) = \mathbf{I}_1(\vartheta_0) + \mathbf{I}_2(\vartheta_0)$$

Denote

$$G_3(\vartheta, \vartheta_0) = \int_0^T \frac{[f(\vartheta, t)\, b(\vartheta, t) - f(\vartheta_0, t)\, b(\vartheta_0, t)]^2}{2f(\vartheta, t)\, b(\vartheta, t)\ \sigma(t)} \mathrm{d}t$$
$$+ \int_0^T \frac{[a(\vartheta, t) - a(\vartheta_0, t) + \ell(\vartheta, t) - \ell(\vartheta_0, t)]^2}{f(\vartheta, t)^2\, b(\vartheta, t)^2} f(\vartheta_0, t)^2\, y_t(\vartheta_0)^2\, \mathrm{d}t.$$

Theorem

Let the conditions of two preceding theorems with Fisher information $\mathbf{I}_3(\vartheta_0)$ *and function* $G_3(\vartheta, \vartheta_0)$ *hold. Then the MLE* $\hat{\vartheta}_\varepsilon$ *and BE* $\tilde{\vartheta}_\varepsilon$ *are uniformly on compacts consistent, asymptotically normal*

$$\varepsilon^{-1/3}\left(\hat{\vartheta}_\varepsilon - \vartheta_0\right) \Longrightarrow \hat{\zeta}(\vartheta_0) \sim \mathcal{N}\left(0, \mathbf{I}_3(\vartheta_0)^{-1}\right),$$
$$\varepsilon^{-1/3}\left(\tilde{\vartheta}_\varepsilon - \vartheta_0\right) \Longrightarrow \hat{\zeta}(\vartheta_0),$$

the moments converge and both estimators are asymptotically efficient.

Case B. $\left|\dot{f}(\vartheta,0)\right| > \kappa > 0$, d=1.

Conditions $\mathscr{G}$.

- $\mathscr{G}_1$. *The functions* $f(\vartheta,0), \dot{f}(\vartheta,0), a(\vartheta,0), b(\vartheta,0)$ *are positive and separated from zero.*
- $\mathscr{G}_2$. $f(\vartheta,t), a(\vartheta,t), b(\vartheta,t) \in \mathcal{C}^2_{\vartheta}$.
- $\mathscr{G}_3$. $f(\vartheta,t), \dot{f}(\vartheta,t), a(\vartheta,t), b(\vartheta,t), \sigma(t) \in \mathcal{C}^1_t$.

The Fisher information function

$$\mathrm{I}_0(\vartheta) = \frac{y_0^2\,\dot{f}(\vartheta,0)^2}{f(\vartheta,0)\,b(\vartheta,0)\,\sigma(0)}, \qquad \vartheta \in \Theta = (\alpha,\beta).$$

Theorem

Let the conditions $\mathscr{G}$ hold. Then the MLE and BE are uniformly consistent, asymptotically normal

$$\varepsilon^{-\frac{1+\delta}{2}}\left(\hat{\vartheta}_{\varepsilon}-\vartheta_0\right)\Longrightarrow\zeta^{\circ}\left(\vartheta_0\right)\sim\mathcal{N}\left(0,\mathrm{I}_0\left(\vartheta_0\right)^{-1}\right),$$

$$\varepsilon^{-\frac{1+\delta}{2}}\left(\tilde{\vartheta}_{\varepsilon}-\vartheta_0\right)\Longrightarrow\zeta^{\circ}\left(\vartheta_0\right),$$

the moments converge and the both estimators are asymptotically efficient.

1 Hidden Markov Processes

2 MDE.

3 MLE and BE.

4 One-step MLE

5 Adaptive filters.

One-step MLE-processes

Our goal is to construct an adaptive Kalman filter which has relatively simple structure and is asymptotically optimal. To do this we need One-step MLE-process which is a modification of the well known Le Cam's One-step MLE. Suppose that we have a consistent estimator $\bar{\vartheta}_\varepsilon$ of ϑ. Then the Le Cam's One-step MLE $\vartheta^\star_\varepsilon$ is like

$$\vartheta^\star_\varepsilon = \bar{\vartheta}_\varepsilon + \mathbf{I}\left(\bar{\vartheta}_\varepsilon\right)^{-1} \int_0^T \frac{\dot{M}(\bar{\vartheta}_\varepsilon, s)}{\sigma\left(s\right)^2} \left[\mathrm{d}X_s - M(\bar{\vartheta}_\varepsilon, s)\mathrm{d}s\right].$$

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

As the estimator $\bar{\vartheta}_\varepsilon$ depends on the same observations the stochastic Itô integral is not defined. Another point: the estimator $\vartheta^\star_\varepsilon$ depends on the whole observations $X^T = (X_t, 0 \le t \le T)$ and can not be used at the moment t for approximation of $m\left(\vartheta_0, t\right)$.

The proposed modifications are

1. Introduce a *learning interval* $[0, \tau]$, where $\tau \in (0, T)$ and construct a preliminary estimator $\bar{\vartheta}_{\tau,\varepsilon}$ by observations X^{τ}. Sometimes $\tau = \tau_{\varepsilon} \to 0$.
2. Define Fisher information function $\mathrm{I}_{\tau}^{t}(\vartheta), t \in (\tau, T]$, which corresponds to the observations on the time interval $[\tau, t]$.
3. Set the low and upper limits of the integral as τ and t respectively. Here $t \in (\tau, T]$.

Then we obtain the One-step MLE-process $\vartheta^{\star}_{t,\varepsilon}, \tau < t \leq T$

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

This estimator has a **recursive representation**. Let us write the matrix

$$\mathbf{I}_{\tau}^{t}\left(\vartheta\right)=\int_{\tau}^{t}\mathrm{P}\left(\vartheta,s\right)\mathrm{d}s$$

with the corresponding matrix $\mathrm{P}\left(\vartheta,s\right)$. Then

$$\mathrm{d}\vartheta_{t,\varepsilon}^{\star}=\mathbf{I}_{\tau}^{t}\left(\bar{\vartheta}_{\tau,\varepsilon}\right)^{-1}\mathrm{P}\left(\bar{\vartheta}_{\tau,\varepsilon},t\right)\left[\bar{\vartheta}_{\tau,\varepsilon}-\vartheta_{t,\varepsilon}^{\star}\right]\mathrm{d}t$$
$$+\mathbf{I}_{\tau}^{t}\left(\bar{\vartheta}_{\tau,\varepsilon}\right)^{-1}\frac{\dot{M}(\bar{\vartheta}_{\tau,\varepsilon},t)}{\sigma\left(t\right)^{2}}\left[\mathrm{d}X_{t}-M(\bar{\vartheta}_{\tau,\varepsilon},t)\mathrm{d}t\right],\quad\tau<t\leq T$$

with the initial value $\vartheta_{\tau,\varepsilon}^{\star}=\bar{\vartheta}_{\tau,\varepsilon}$.

Therefore the realization of the One-step MLE-process can be easily done using numerical methods.

Case A, $\dot{f}(\vartheta, 0) = 0$**. Fixed learning interval.** $\delta \in (1/3, 1)$**.**

We are given the partially observed system with $\psi_\varepsilon = \varepsilon^\delta$, $\delta \in (0, 1)$. The learning interval is $[0, \tau]$ and the preliminary MDE is

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

Note that if the conditions $\mathscr{B}$ hold then by the Theorem the MDE is consistent and asymptotically normal

$$\psi_\varepsilon^{-1}\left(\check{\vartheta}_{\tau,\varepsilon} - \vartheta_0\right) \Longrightarrow \check{\zeta}(\vartheta_0) \sim \mathcal{N}\left(0, \check{D}(\vartheta_0)\right).$$

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

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

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

$$A^*(\vartheta, t) = a^*(\vartheta, t) - \frac{\psi_\varepsilon}{\varepsilon}\tilde{B}(\vartheta, t), \qquad \tilde{B}(\vartheta, t) = \frac{\gamma_*(\vartheta, t)\, f(\vartheta, t)^2}{\sigma(t)^2},$$

$$a^*(\vartheta, t) = a(\vartheta, t) + \ell'_t(\vartheta, t), \qquad \ell(\vartheta, t) = \ln f(\vartheta, t).$$

Denote

$$F(\vartheta,t)=\frac{[\dot{a}(\vartheta,t)+\dot{\ell}(\vartheta,t)]y_t(\vartheta)}{b(\vartheta,t)\sigma(t)},\qquad Q(\vartheta,t)=F(\vartheta,t)\sigma(t)$$

and Fisher information function

$$\mathbf{I}_0^t(\vartheta)=\int_0^t\frac{[\dot{a}(\vartheta,s)+\dot{\ell}(\vartheta,s)][\dot{a}(\vartheta,s)+\dot{\ell}(\vartheta,s)]^\top y_s(\vartheta)^2}{b(\vartheta,s)^2}\mathrm{d}s.$$

We need *Conditions* $\mathscr{F}^\star$.

$\mathscr{F}_1^\star$. $\dot{a}(\vartheta,t),\dot{\ell}(\vartheta,t),b(\vartheta,t),\sigma(t)\in\mathcal{C}_t^1$.

$\mathscr{F}_2^\star$. *The Fisher information function* $\mathbf{I}_0^t(\vartheta),0<\tau<t\leq T$ *is uniformly on* $\vartheta\in\Theta$ *non degenerate for all* $t\in(\tau,T]$.

Introduce the One-step MLE-process $\vartheta^{\star}_{t,\varepsilon}, \tau \le t \le T$

$$\vartheta^{\star}_{t,\varepsilon} = \check{\vartheta}_{\tau,\varepsilon} + \mathrm{I}^{t}_{0}(\check{\vartheta}_{\tau,\varepsilon})^{-1}\frac{\psi_{\varepsilon}}{\varepsilon}\Big\{F(\check{\vartheta}_{\tau,\varepsilon},\tau)X_{\tau} - \int_{0}^{\tau} X_{s}\,F'_{s}(\check{\vartheta}_{\tau,\varepsilon},s)\mathrm{d}s$$
$$- \int_{0}^{\tau} F(\check{\vartheta}_{\tau,\varepsilon},s)M(\check{\vartheta}_{\tau,\varepsilon},s)\mathrm{d}s + \int_{\tau}^{t} F(\check{\vartheta}_{\tau,\varepsilon},s)\,\mathrm{d}X_{s}\Big\}.$$

Theorem

Let the conditions $\mathscr{B}$, $\mathscr{F}_{0-3}$, $\mathscr{F}^{\star}$ hold and $\delta \in (1/3, 1)$. Then the estimator process $\vartheta^{\star}_{t,\varepsilon}, \tau < t \le T$ is uniformly consistent, asymptotically normal

$$\psi_{\varepsilon}^{-1}(\vartheta^{\star}_{t,\varepsilon} - \vartheta_{0}) \Longrightarrow \zeta_{t}(\vartheta_{0}) \sim \mathcal{N}\left(0, \mathbf{I}^{t}_{0}(\vartheta_{0})^{-1}\right),$$

the moments converge and this estimator is asymptotically efficient.

Case A, $\dot{f}(\vartheta, 0) = 0$**. Fixed learning interval.** $\delta \in (0, 1/3)$**.**

Remind that in this case $\varphi_\varepsilon = \varepsilon^{\frac{1-\delta}{2}}$

$$\frac{M(\vartheta_0 + \varphi_\varepsilon u, t) - M(\vartheta_0, t)}{\varepsilon \sigma(t)} = \frac{u^\top \dot{S}(\vartheta_0, t)}{\sqrt{2B(\vartheta_0, t)}\, \sigma(t)} \xi_{t,\varepsilon} (1 + O(\varphi_\varepsilon))$$

where $S(\vartheta, t) = f(\vartheta, t)\, b(\vartheta, t)$ and $B(\vartheta, t) = S(\vartheta, t)/\sigma(t)$. The Fisher information matrix-function is

$$\mathrm{I}_\tau^t(\vartheta) = \int_\tau^t \frac{\dot{S}(\vartheta, s)\, \dot{S}(\vartheta, s)^\top}{2S(\vartheta, s)\, \sigma(s)}\, \mathrm{d}s, \qquad \tau < t \le T.$$

We have two preliminary MDEs. The first one $\check{\vartheta}_{\tau,\varepsilon}$ with fixed learning interval $[0,\tau]$ and the second $\check{\vartheta}_{\tau_\varepsilon}$ (one-dimensional ϑ) - with the vanishing interval $[0,\tau_\varepsilon]$, $\tau_\varepsilon = [\ln \varepsilon^{-1}]^{-1} \to 0$. The rates of convergence of MDEs are ε^δ and $\varepsilon^\delta \ln \varepsilon^{-1}$ respectively. Suppose as well that the conditions $\mathscr{B}$ hold. Then by Theorem we have

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

It is shown that if we define the One-step MLE-process by the usual relation then the last term has the expression

$$\left[\psi_\varepsilon^{-1}\left(\check{\vartheta}_{\tau,\varepsilon} - \vartheta_0\right)\right]^2 \varepsilon^{\frac{5\delta-1}{2}}.$$

It tends to 0 for the values $\delta > 1/5$. That is why we consider the cases $\delta \in (1/5, 1/3)$ and $\delta \in (0, 1/5)$ separately.

Let $\delta \in (1/5, 1/3)$. The One-step MLE-process is

$$\vartheta^{\star}_{t,\varepsilon} = \check{\vartheta}_{\tau,\varepsilon} + \mathrm{I}^{t}_{\tau}(\check{\vartheta}_{\tau,\varepsilon})^{-1} \int_{\tau}^{t} \frac{\dot{M}(\check{\vartheta}_{\tau,\varepsilon}, s)}{\varepsilon\, \psi_{\varepsilon}\, \sigma\,(s)^{2}} \left[\mathrm{d}X_{s} - M(\check{\vartheta}_{\tau,\varepsilon}, s)\mathrm{d}s \right].$$

Conditions $\mathscr{F}^{\star}$.

$\mathscr{F}^{\star}_{3}$. $\dot{b}\,(\vartheta, t)\,, \dot{f}\,(\vartheta, t)\,, f\,(\vartheta, t)\,, \sigma\,(t) \in \mathcal{C}^{1}_{t}$.

$\mathscr{F}^{\star}_{4}$. *Fisher information* $\mathbf{I}^{t}_{\tau}\,(\vartheta)\,, \tau < t \leq T$ *is unif. positive.*

Theorem

Let the conditions $\mathscr{B}$, $\mathscr{F}_{1-3}$ *and* $\mathscr{F}^{\star}_{3-4}$ *hold,* $\varphi_{\varepsilon} = \varepsilon^{\frac{1-\delta}{2}}$ *and* $\delta \in (1/5, 1/3)$. *Then the estimator process* $\vartheta^{\star}_{t,\varepsilon}, \tau < t \leq T$ *is uniformly consistent, asymptotically normal*

$$\varphi^{-1}_{\varepsilon}\left(\vartheta^{\star}_{t,\varepsilon} - \vartheta_{0}\right) \Longrightarrow \mathcal{N}\left(0, \mathbf{I}^{t}_{\tau}\,(\vartheta_{0})^{-1}\right),$$

the moments converge and this estimator is asymptotically efficient.

Case **B**, $\left|\dot{f}(\vartheta,0)\right| \geq 0$. $\delta \in (0,1)$

Suppose now that $d=1$ and the MDE $\check{\vartheta}_{\tau_\varepsilon}$ is constructed by observations on the vanishing interval $[0,\tau_\varepsilon]$, where $\tau_\varepsilon = \varepsilon^{1-\delta}$. Then under conditions $\mathscr{D}^\circ$, $\phi_\varepsilon = \varphi_\varepsilon = \sqrt{\varepsilon\psi_\varepsilon}$ and

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

The Fisher information is

$$\mathrm{I}_0\left(\vartheta\right) = \frac{y_0^2\ \dot{f}\left(\vartheta,0\right)^2}{f\left(\vartheta,0\right) b\left(\vartheta,0\right) \sigma\left(0\right)}.$$

The score-function for any $t > t_0$ is

$$\frac{\partial L\left(\vartheta, X^t\right)}{\partial \vartheta} = \frac{y_0\, \dot{f}\left(\vartheta,0\right)}{\varepsilon^2\, \sigma\left(0\right)^2} \int_0^{t_0} \frac{\left[\mathrm{d}X_s - M\left(\vartheta,s\right)\mathrm{d}s\right]}{\cosh\left(\frac{\psi_\varepsilon}{\varepsilon}\Gamma\left(\vartheta\right)s\right)}\left(1+o\left(1\right)\right).$$

Here $\Gamma\left(\vartheta\right) = f\left(\vartheta,0\right) b\left(\vartheta,0\right) \sigma\left(0\right)^{-1}$.

We cannot substitute $\check{\vartheta}_{\tau_\varepsilon}$ directly in this stochastic integral. Denote $H(x) = [\cosh(x)]^{-1}$ and write

$$\int_0^{t_0} H\left(\frac{\psi_\varepsilon}{\varepsilon}\Gamma(\vartheta)\, s\right) \mathrm{d}X_s = H\left(\frac{\psi_\varepsilon}{\varepsilon}\Gamma(\vartheta)\, t_0\right) X_{t_0} - \frac{\psi_\varepsilon}{\varepsilon}\Gamma(\vartheta) \int_0^{t_0} H'\left(\frac{\psi_\varepsilon}{\varepsilon}\Gamma(\vartheta)\, s\right) X_s \mathrm{d}s.$$

Introduce the function $\mathrm{K}(\vartheta)$ by the relation

$$\mathrm{I}_0(\vartheta)^{-1} \frac{y_0 \dot{f}(\vartheta, 0)}{\sigma(0)^2} = \frac{f(\vartheta, 0)\, b(\vartheta, 0)}{y_0 \dot{f}(\vartheta, 0)\, \sigma(0)} = \mathrm{K}(\vartheta).$$

The One-step MLE we define as follows

$$\vartheta^{\star}_{\varepsilon} = \check{\vartheta}_{\tau_\varepsilon} + \mathrm{K}(\check{\vartheta}_{\tau_\varepsilon})\,\frac{\psi_\varepsilon}{\varepsilon}\,\Bigg[\frac{X_{t_0}}{\cosh(\frac{\psi_\varepsilon}{\varepsilon}\Gamma(\check{\vartheta}_{\tau_\varepsilon})t_0)} - \frac{\psi_\varepsilon}{\varepsilon}\Gamma(\check{\vartheta}_{\tau_\varepsilon})\int_0^{t_0} H'(\frac{\psi_\varepsilon}{\varepsilon}\Gamma(\check{\vartheta}_{\tau_\varepsilon})s)X_s\,\mathrm{d}s\Bigg] - \mathrm{K}(\check{\vartheta}_{\tau_\varepsilon})\,\frac{\psi_\varepsilon}{\varepsilon}\int_0^{t_0}\frac{M(\check{\vartheta}_{\tau_\varepsilon},s)}{\cosh(\frac{\psi_\varepsilon}{\varepsilon}\Gamma(\check{\vartheta}_{\tau_\varepsilon})s)}\,\mathrm{d}s.$$

Theorem

The One-step MLE $\vartheta^{\star}_{\varepsilon}$ is consistent and asymptotically normal

$$\varepsilon^{-\frac{1+\delta}{2}}(\vartheta^{\star}_{\varepsilon} - \vartheta_0) \Longrightarrow \mathcal{N}(0, \mathrm{I}_0(\vartheta_0)^{-1}).$$

1 Hidden Markov Processes

2 MDE.

3 MLE and BE.

4 One-step MLE

5 Adaptive filters.

Adaptive filters.Case A, $\delta \in (1/3, 1)$.

Suppose that the learning interval is $[0, \tau]$, $0 < \tau < T$ and the MDE $\check{\vartheta}_{\tau,\varepsilon}$. This estimator is asymptotically normal with the rate ψ_ε. This allow us to admit the values $\delta \in (1/3, 1)$ for the construction of the One-step MLE-process $\vartheta^\star_{t,\varepsilon}, \tau < t \leq T$. Recall that

$$\frac{\vartheta^\star_{t,\varepsilon} - \vartheta_0}{\psi_\varepsilon} \Longrightarrow \eta^\star_t(\vartheta_0) = \mathrm{I}^t_0(\vartheta_0)^{-1} \int_0^t \frac{[\dot{a}(\vartheta_0, s) + \dot{\ell}(\vartheta_0, s)] y_s(\vartheta_0)}{b(\vartheta_0, s)} \mathrm{d}w(s)$$

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 $\tau_* \in (\tau, T)$.

Proposition

Let the conditions $\mathscr{B}$, $\mathscr{F}_{1-3}$ *and* $\mathscr{F}^{\star}$ *hold. Then for* $\delta \in (1/3, 1)$

$$\frac{m^{\star}_{t,\varepsilon} - m\left(\vartheta_0, t\right)}{\psi_{\varepsilon}} \Longrightarrow -\frac{\dot{f}\left(\vartheta_0, t\right)\ y_t\left(\vartheta_0\right)}{f\left(\vartheta_0, t\right)}\ \eta^{\star}_t\left(\vartheta_0\right).$$

It is possible to prove the asymptotic efficiency of the adaptive filter with the help of the following lower bound: Let the conditions $\mathscr{D}$ hold then for any estimator $\bar{m}_{t,\varepsilon}$ and any $\vartheta_0 \in \Theta$ we have

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

where

$$\mathcal{S}^{\star}_t\left(\vartheta_0\right)^2 = \frac{\dot{f}\left(\vartheta_0, t\right)^2\ y_t\left(\vartheta_0\right)^2}{f\left(\vartheta_0, t\right)^2}\ \mathrm{I}^t_0\left(\vartheta_0\right)^{-1}.$$

Adaptive filter with "=" is called asymptotically optimal.

B Case. Vanishing interval. $\delta \in (0,1)$.

The One-step MLE was given above and it is asymptotically normal

$$\varphi_{\varepsilon}^{-1}\left(\vartheta_{\varepsilon}^{\star}-\vartheta_{0}\right) \Longrightarrow \eta^{\star}\left(\vartheta_{0}\right)=\mathrm{I}_{0}\left(\vartheta_{0}\right)^{-1/2} \int_{0}^{\infty} \frac{\mathrm{d}w\left(x\right)}{\cosh\left(x\right)}.$$

Here $w\left(x\right), x \geq 0$ is a Wiener process.

The adaptive filter is defined by the equation

$$\mathrm{d}m_{t,\varepsilon}^{\star}=\left[a(\vartheta_{\varepsilon}^{\star},t)-\frac{\psi_{\varepsilon}}{\varepsilon}\frac{b\left(\vartheta_{\varepsilon}^{\star},t\right)f\left(\vartheta_{\varepsilon}^{\star},t\right)}{\sigma(t)}\right]m_{t,\varepsilon}^{\star}\mathrm{d}t+\frac{\psi_{\varepsilon}}{\varepsilon}\frac{b\left(\vartheta_{\varepsilon}^{\star},t\right)}{\sigma(t)}\,\mathrm{d}X_{t}$$

for $t_0 \leq t \leq T$, where t_0 is any (small) fixed value. Moreover, in this case we can write directly $m_{t,\varepsilon}^{\star}=m\left(\vartheta_{\varepsilon}^{\star},t\right)$ and the error of approximation we obtain from this equality

$$\varepsilon^{-\frac{1+\delta}{2}}\left(m_{t,\varepsilon}^{\star}-m\left(\vartheta_{0},t\right)\right) \Longrightarrow -\frac{\dot{f}\left(\vartheta_{0},t\right)\,y_{t}\left(\vartheta_{0}\right)}{f\left(\vartheta_{0},t\right)}\,\eta^{\star}\left(\vartheta_{0}\right)$$

Thank you for your attention!