On Adaptive Kalman Filtration.

Yury A. Kutoyants

Le Mans University, France

ICIAM 2023 TOKYO

25 August, 2023

1 Hidden Markov Processes

1.1 Model of observations.

We are given a couple of equations

$$\mathrm{d}X_t = f(\vartheta, t)\, Y_t \mathrm{d}t + \sigma(t)\, \mathrm{d}W_t, \qquad X_0, \quad 0 \le t \le T,$$

$$\mathrm{d}Y_t = a(\vartheta, t)\, Y_t \mathrm{d}t + b(\vartheta, t)\, \mathrm{d}V_t, \qquad Y_0, \quad 0 \le t \le T,$$

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 O-U 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 and the parameter $\vartheta \in \Theta \subset \mathcal{R}^d$ is unknown.

The conditional expectation $m(\vartheta,t) = \mathbf{E}_\vartheta \left(Y_t | X_s, 0 \le s \le t\right)$ 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

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

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

with initial values $m_0 = \mathbf{E}_\vartheta \left(Y_0 | X_0\right)$ and $\gamma_0 = \mathbf{E}_\vartheta \left(Y_0 - m(\vartheta,0)\right)^2$. Recall that $m(\vartheta,t)$ is an optimal estimator of Y_t.

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

We need a good estimator of ϑ, which depends on $X_s, 0 \le s \le t$ to use it for construction of $m_t^\star, 0 < t \le T$.

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

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

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 $\left\{m\left(\vartheta,t\right),\ 0\leq t\leq T\right\},\ \vartheta\in\Theta$ and $\left\{\gamma\left(\vartheta,t\right),\ 0\leq t\leq T\right\},\ \vartheta\in\Theta$ defined by the K-B equations.

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 $m(\vartheta, t)$ we propose the following program:

1. *Find a preliminary estimator $\bar\vartheta_\tau$ on relatively small interval of observations $[0, \tau]$.*
2. *Using $\bar\vartheta_\tau$ realize One-step MLE-process $\vartheta_t^\star, \tau < t \leq T$*
3. *As approximation of $m(\vartheta, t)$ 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 $m_t^\star - m(\vartheta, t)$.*

We apply this construction to 5 different models of observations.

1.2 HMP with small noises in both equations.

Consider the linear two-dimensional partially observed system

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

Asymptotic $\varepsilon \to 0$. The preliminary estimator is $(x_t(\vartheta) = X_t|_{\varepsilon=0})$

$$
\check{\vartheta}_{\tau_\varepsilon} = \arg \inf_{\vartheta \in \Theta} \int_0^{\tau_\varepsilon} \left| X_t - x_t(\vartheta) \right|^2 \mathrm{d}t, \qquad \tau_\varepsilon \to 0
$$

YuK, (1994) Identification of Dynamical Systems with Small Noise. Kluwer Academic Publisher, Dordrecht.

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

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

$$\dot S_*\left(\vartheta,\vartheta_0,t\right)=\frac{\partial S_*\left(\vartheta,\vartheta_0,t\right)}{\partial\vartheta}=\dot f\left(\vartheta,t\right)y_t\left(\vartheta,\vartheta_0\right)+f\left(\vartheta,t\right)\dot y_t\left(\vartheta,\vartheta_0\right),$$

$$\mathbf{I}\left(\vartheta_0\right)=\int_0^T\frac{\dot S_*\left(\vartheta_0,\vartheta_0,t\right)\dot S_*\left(\vartheta_0,\vartheta_0,t\right)^\top}{\sigma\left(t\right)^2}{\rm d}t.$$

The $d\times d$ matrix $\mathbf{I}\left(\vartheta_0\right),\vartheta_0\in\Theta$ is the Fisher information. The MLE and BE are uniformly on compacts $\mathbb{K}\subset\Theta$ consistent, asymptotically normal

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

One-step MLE-process $\vartheta^\star_{t,\varepsilon}, \tau_\varepsilon < t \leq T$ are defined by the relations

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

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

Here $M\left(\vartheta, s\right) = f\left(\vartheta, s\right) m\left(\vartheta, s\right)$ and

$$\mathbf{I}^t_{\tau_\varepsilon}\left(\vartheta_0\right) = \int_{\tau_\varepsilon}^t \frac{\dot S_*\left(\vartheta_0, \vartheta_0, s\right)\dot S_*\left(\vartheta_0, \vartheta_0, s\right)^\top}{\sigma\left(s\right)^2}{\rm d}s$$

This estimator-process is as. normal

$$\frac{\vartheta^\star_{t,\varepsilon} - \vartheta_0}{\varepsilon} \Longrightarrow \zeta \sim \mathcal{N}\left(0, \mathbf{I}^t_0\left(\vartheta_0\right)^{-1}\right),$$

This estimator is used for adaptive filtration.

1.3 HMP with low noise observations

Consider the linear two-dimensional partially observed system

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

$$\mathrm{d}Y_t = a(\vartheta, t)\, Y_t \mathrm{d}t + b(\vartheta, t)\, \mathrm{d}V_t, \quad Y_0,$$

Asymptotic $\varepsilon \to 0$. Below $\tau_\varepsilon = \varepsilon^{1/6}, t_{i+1} - t_i = \varepsilon^{2/3}, N_\varepsilon = \varepsilon^{-2/3}$

$$\Psi_{\tau_\varepsilon,\varepsilon} = \sum_{i=0}^{N_\varepsilon - 1} \left(\frac{X_{t_{i+1}+\varepsilon} - X_{t_{i+1}}}{\varepsilon} - \frac{X_{t_i+\varepsilon} - X_{t_i}}{\varepsilon} \right)^2,$$

$$\int_0^{\tau_\varepsilon} f\left(\check\vartheta_{\tau_\varepsilon}, t\right)^2 b\left(\check\vartheta_{\tau_\varepsilon}, t\right)^2 \mathrm{d}t = \Psi_{\tau_\varepsilon,\varepsilon}$$

YuK (2019) "On parameter estimation of hidden Ornstein -Uhlenbeck process", J. Multivariate Analysis. *169, 1, 248-263.*

YuK (2022) "Volatility estimation of hidden Markov process and adaptive filtration", arXiv:2010.07603, submitted

Fisher information matrix is

$$\mathbf{I}\left(\vartheta\right)=\int_{0}^{T}\frac{S\left(\vartheta,t\right)}{2\sigma\left(t\right)}\left[\frac{\partial}{\partial\vartheta}\ln S\left(\vartheta,t\right)\right]\left[\frac{\partial}{\partial\vartheta}\ln S\left(\vartheta,t\right)\right]^{\top}\mathrm{d}t.$$

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

The MLE $\hat\vartheta_\varepsilon$ and BE $\tilde\vartheta_\varepsilon$ are consistent, and asymptotically normal, i.e.,

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

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

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

This estimator is used for adaptive filtration.

1.4 Extended Kalman Filter

The system is

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

The functions $f\left(\cdot\right), \sigma\left(\cdot\right), a\left(\cdot\right), b\left(\cdot\right)$ are known and positive,
$f'_y\left(\vartheta, t, y\right) \geq \kappa > 0$

The equations of *Extended Kalman Filter* (EKF) are

$$\mathrm{d}m_t(\vartheta) = \left[a(\vartheta, t, m_t(\vartheta)) - \frac{f'_y(\vartheta, t, m_t(\vartheta))\, f(\vartheta, t, m_t(\vartheta))\, \gamma(\vartheta, t)}{\varepsilon^2 \sigma(t)^2}\right] \mathrm{d}t$$

$$+ \frac{f'_y(\vartheta, t, m_t(\vartheta))\, \gamma_*(\vartheta, t)}{\varepsilon^2 \sigma(t)^2}\, \mathrm{d}X_t, \qquad m_0(\vartheta) = y_0,$$

and

$$\frac{\partial \gamma(\vartheta, t)}{\partial t} = 2a'_y(\vartheta, t, m_t(\vartheta))\, \gamma(\vartheta, t) - \frac{f'_y(\vartheta, t, m_t(\vartheta))^2\, \gamma(\vartheta, t)^2}{\varepsilon^2 \sigma(t)^2}$$

$$+ b(\vartheta, t, m_t(\vartheta))^2, \qquad \gamma(\vartheta, 0) = 0.$$

If the observed system is linear then these equations coincide with the usual Kalman filter.

We have the same problem of adaptive filtration and almost the same MME $\check{\vartheta}_{\tau,\varepsilon}$ based on the statistic

$$\Psi_{\tau,\varepsilon} = \sum_{i=0}^{N_\varepsilon - 1} \left(\frac{X_{t_{i+1}+\varepsilon} - X_{t_{i+1}}}{\varepsilon} - \frac{X_{t_i+\varepsilon} - X_{t_i}}{\varepsilon} \right)^2 \longrightarrow \Psi_\tau \left(\vartheta_0\right),$$

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

$$\int_0^\tau f\left(\check{\vartheta}_{\tau,\varepsilon}, t, m_t(\check{\vartheta}_{\tau,\varepsilon})\right)^2 b\left(\check{\vartheta}_{\tau,\varepsilon}, t, m_t(\check{\vartheta}_{\tau,\varepsilon}\right)^2 \mathrm{d}t = \Psi_{\tau,\varepsilon}$$

where $t_{i+1} - t_i = \varepsilon^{2/3}, N_\varepsilon = \tau\varepsilon^{-2/3}$. Under regularity conditions

$$\varepsilon^{-1/3}(\check{\vartheta}_{\tau,\varepsilon} - \vartheta_0)$$

is tight.

We have the following amusing relations for $t_0 \leq t \leq T$, :

$$\frac{\gamma\left(\vartheta,t\right)}{\varepsilon} \longrightarrow \gamma_*\left(\vartheta,t\right) = \frac{b\left(\vartheta,t\right)\sigma\left(t\right)}{f'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)},$$

$$f\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right) = f\left(\vartheta_0,t,Y_t\right),$$

$$m_t^\circ\left(\vartheta_0\right) = Y_t,$$

$$\dot m_t^\circ\left(\vartheta\right) = -\frac{\dot f_\vartheta\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}, \qquad \dot m_t^\circ\left(\vartheta_0\right) = -\frac{\dot f_\vartheta\left(\vartheta_0,t,Y_t\right)}{f_y'\left(\vartheta_0,t,Y_t\right)}.$$

This means that the "wrong" filter $m_t\left(\vartheta_0\right)$ converges to the true process $m_t\left(\vartheta_0\right) \to m_t^\circ\left(\vartheta_0\right) = Y_t$.

Denote

$$\Gamma_{t,\varepsilon}\left(\vartheta\right)=\sqrt{\frac{b\left(\vartheta,t\right)\sigma\left(t\right)}{f'_y\left(\vartheta,t,m_t\left(\vartheta\right)\right)}},\qquad \Gamma_t^\circ\left(\vartheta\right)=\sqrt{\frac{b\left(\vartheta,t\right)\sigma\left(t\right)}{f'_y\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}},$$

$$\Gamma_t^\circ\left(\vartheta_0\right)=\sqrt{\frac{b\left(\vartheta_0,t\right)\sigma\left(t\right)}{f'_y\left(\vartheta_0,t,Y_t\right)}},\qquad \Delta_t^\circ\left(\vartheta_0\right)=\Gamma_t^\circ\left(\vartheta_0\right)\left[\xi_{t,\varepsilon}-\eta_{t,\varepsilon}\right]$$

$$\Delta_{t,\varepsilon}\left(\vartheta\right)=\Gamma_{t,\varepsilon}\left(\vartheta\right)\left[\xi_{t,\varepsilon}-\frac{b\left(\vartheta_0,t\right)f'_y\left(\vartheta_0,t,Y_t\right)}{b\left(\vartheta,t\right)f'_y\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\;\eta_{t,\varepsilon}\right].$$

We have

$$m_t\left(\vartheta\right)=m_t^\circ\left(\vartheta\right)+\Delta_{t,\varepsilon}\left(\vartheta\right)\;\sqrt{\varepsilon}+O\left(\varepsilon\right)$$

and

$$\dot m_t\left(\vartheta\right)=-\frac{\dot f_\vartheta\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{f'_y\left(\vartheta,t,m_t\left(\vartheta\right)\right)}-\frac{\dot b\left(\vartheta,t\right)}{b\left(\vartheta,t\right)}\,\Delta_{t,\varepsilon}\left(\vartheta\right)\sqrt{\varepsilon}+\left[\ldots\right]\xi_{t,\varepsilon}\sqrt{\varepsilon}+O\left(\varepsilon\right)$$

This allows us to construct a p-One-step process and adaptive EKF.

1.5 Hidden Telegraph process.

The observations are

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

where $Y_t, 0 \le t \le T$ is *telegraph process* and $W_t, 0 \le t \le T$ is an independent of $Y_t, 0 \le t \le T$ Wiener process. Recall that $Y_t, t \ge 0$ is a memory-less continuous-time stochastic Markov process that shows two distinct values $y_1 = a$ and $y_2 = b$. Parameter $\vartheta = (\lambda, \mu)$. This process can be described by the transition rate matrix

$$\begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}.$$

The stochastic process $Y_t, t \geq 0$ admits the representation

$$\mathrm{d}X_t = m(\vartheta, t)\ \mathrm{d}t + \mathrm{d}\bar{W}_t, \qquad X_0, \qquad 0 \leq t \leq T,$$

where $m(\vartheta, t) = \mathbf{E}_\vartheta \left(Y_t \middle| \mathfrak{F}_t^X \right)$ is the conditional expectation. Let us denote $\pi(\vartheta, t) = \mathbf{P}_\vartheta \left(Y_t = a \middle| \mathfrak{F}_t^X \right), \mathbf{P}_\vartheta \left(Y_t = b \middle| \mathfrak{F}_t^X \right) = 1 - \pi(\vartheta, t)$. Then $m(\vartheta, t) = b + (a-b)\pi(\vartheta, t)$.

The random process $\pi(\vartheta, \cdot)$ satisfies the following equation

$$\begin{aligned}\mathrm{d}\pi(\vartheta, t) = [&\mu - (\lambda + \mu)\pi(\vartheta, t) \\ &+\pi(\vartheta, t)(1 - \pi(\vartheta, t))(b-a)(b + (a-b)\pi(\vartheta, t))]\,\mathrm{d}t \\ &+ \pi(\vartheta, t)(1 - \pi(\vartheta, t))(a-b)\ \mathrm{d}X_t.\end{aligned}$$

Chigansky, P. (2009). "Maximum likelihood estimator for hidden Markov models in continuous time". SISP. 12, 139-163.

Khasminskii, R. Z. and YuK (2018) "On parameter estimation of hidden telegraph process". Bernoulli, 24, 3, 2064-2090.

Having a preliminary MME $\check{\vartheta}_{\tau_\delta}$ we propose One-step MLE-process as follows $\left(\tau_\delta = T^\delta, \delta \in \left(\frac{1}{2}, 1\right)\right)$

$$\vartheta^\star_{t,T} = \check{\vartheta}_{\tau_\delta} + t^{-1}\mathbb{I}_t(\check{\vartheta}_{\tau_\delta})^{-1} \int_{\tau_\delta}^{t} \dot{m}(\check{\vartheta}_{\tau_\delta}, s) \left[\mathrm{d}X_s - m(\check{\vartheta}_{\tau_\delta}, s)\mathrm{d}s \right].$$

Here the vector

$$\dot{m}(\vartheta, s) = (a-b)\,\frac{\partial \pi_\lambda\left(s, \vartheta\right)}{\partial \vartheta} = (a-b)\left(\frac{\partial \pi\left(t, \vartheta\right)}{\partial \lambda}, \frac{\partial \pi\left(t, \vartheta\right)}{\partial \mu}\right)^\top$$

and the empirical Fisher information matrix $\mathbb{I}_t(\vartheta)$ is

$$\mathbb{I}_t(\vartheta) = \frac{1}{t}\int_{\tau_\delta}^{t} \dot{m}(\vartheta, s)\dot{m}(\vartheta, s)^\top \mathrm{d}s \longrightarrow \mathbb{I}(\vartheta)$$

as $t \to \infty$ by the law of large numbers. Here $\mathbb{I}(\vartheta)$ is the Fisher information matrix

$$\mathbb{I}(\vartheta) = (a-b)^2\,\mathbf{E}_\vartheta \frac{\partial \pi\left(s, \vartheta\right)}{\partial \vartheta}\frac{\partial \pi\left(s, \vartheta\right)}{\partial \vartheta}^\top.$$

1.6 Hidden O-U process, Ergodic case

We are given a linear system with constant coefficients

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

Here $f(\vartheta) \neq 0$, $b(\vartheta) \neq 0$, $\sigma^2 > 0$ and $a(\vartheta) > 0$. Asymptotic $T \to \infty$. Examples: $(\mathbf{F}) : f(\vartheta) = \vartheta, \quad a(\vartheta) = a, \quad b(\vartheta) = b$, $(\mathbf{A}) : f(\vartheta) = f, \quad a(\vartheta) = \vartheta, \quad b(\vartheta) = b$, $(\mathbf{F}, \mathbf{B})$ forbidden, $(\mathbf{F}, \mathbf{A}) : f(\vartheta) = \theta_1, \quad a(\vartheta) = \theta_2, \quad b(\vartheta) = b, i.e.\ \vartheta = (\theta_1, \theta_2)$.

YuK (2019) "On parameter estimation of hidden ergodic Ornstein-Uhlenbeck process" Electr. J. of Stat., *13, 4508-4526.*

YuK (2022) "Hidden ergodic Ornstein-Uhlenbeck process and adaptive filter" submitted.

Consider the model $f(\vartheta) = f$,$a(\vartheta) = a$,$b(\vartheta) = b$, and denote

$$R_{1,T} = \frac{1}{T}\sum_{k=1}^{T}\left[X_k - X_{k-1}\right]^2, \quad \Phi_1(\vartheta) = \frac{f^2 b^2}{a^3}\left[e^{-a} - 1 + a\right] + \sigma^2,$$

$$R_{2,T} = \frac{1}{T}\sum_{k=2}^{T}\left[X_k - X_{k-1}\right]\left[X_{k-1} - X_{k-2}\right], \quad R_T = \left(R_{1,T}, R_{2,T}\right)^\top,$$

$$\Phi_2(\vartheta) = \frac{f^2 b^2}{2a^3}\left[1 - e^{-a}\right]^2, \qquad \Phi(\vartheta) = \left(\Phi_1(\vartheta), \Phi_2(\vartheta)\right)^\top,$$

$$K_{1,1}(\vartheta) = \frac{2f^4 b^4}{a^6}\left[e^{-a} - 1 + a\right]^2 + \frac{f^4 b^4 e^{4a}\left[1 - e^{-a}\right]^3}{a^6\left[1 + e^{-a}\right]} + \frac{4f^2\sigma^2 b^2}{a^3}\left[e^{-a} - 1 + a\right] + 2\sigma^4,$$

$$\mathbf{K}(\vartheta) = \begin{pmatrix} K_{1,1}(\vartheta) & K_{1,2}(\vartheta) \\ K_{2,1}(\vartheta) & K_{2,2}(\vartheta) \end{pmatrix}, \qquad \xi_* = \left(\xi_1, \xi_2\right)^\top, .$$

Here $\xi_* \sim \mathcal{N}\left(0, \mathbf{K}(\vartheta_0)\right)$.

Proposition 1. *The statistic R_T has the following properties:*

1. *Uniformly on Θ it converges to $\Phi(\vartheta)$, i.e., for any $\nu > 0$*

$$\lim_{T\to\infty} \sup_{\vartheta\in\Theta} \mathbf{P}_\vartheta \left(\left\|R_T - \Phi(\vartheta)\right\| \geq \nu\right) = 0.$$

2. *Is uniformly on compacts $\mathbb{K} \subset \Theta$ asymptotically normal*

$$\sqrt{T}\left(R_T - \Phi(\vartheta)\right) \Longrightarrow \xi_*,$$

3. *The moments converge: for any $p > 0$ uniformly on $\mathbb{K}$*

$$\lim_{T\to\infty} T^{p/2}\mathbf{E}_\vartheta \left\|R_T - \Phi(\vartheta)\right\|^p = \mathbf{E}_\vartheta \left|\xi_*\right|^p,$$

and there exists a constant $C > 0$ such that

$$\sup_{\vartheta\in\mathbb{K}} T^{p/2}\mathbf{E}_\vartheta \left\|R_T - \Phi(\vartheta)\right\|^p \leq C.$$

1.6.1 Two-dimensional parameter (f, a)

$$\Phi_1(\vartheta) = \frac{f^2 b^2}{a^3}\left[e^{-a} - 1 + a\right] + \sigma^2, \qquad \Phi_2(\vartheta) = \frac{f^2 b^2}{2a^3}\left[1 - e^{-a}\right]^2.$$

Therefore this system can be solved in two steps as follows: the equation

$$\frac{e^{-a_T^*} - 1 + a_T^*}{\left(1 - e^{-a_T^*}\right)^2} = \frac{R_{1,T} - \sigma^2}{2R_{2,T}},$$

$$f_T^* = \frac{2\left(a_T^*\right)^{3/2} R_{2,T}^{1/2}}{b^2\left(1 - e^{-a_T^*}\right)^2}.$$

Having $\left(f_T^*, a_T^*\right)$ we write the One-step process and adaptive filter.

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

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

Change the variables: $t = vT, v \in \left[\varepsilon_T, 1\right], \varepsilon_T = \tau_T T^{-1} = T^{-1+\delta} \rightarrow 0$ and denote $\vartheta^\star_T\left(v\right) = \vartheta^\star_{vT,T}, \varepsilon_T < v \leq 1$.

Then One-step MLE-process $\vartheta^\star_T\left(v\right), \varepsilon_T < v \leq 1$ with $\delta \in (1/2, 1)$ is uniformly on compacts $\mathbb{K} \subset \Theta$ consistent and asymptotically normal

$$\sqrt{T}\left(\vartheta^\star_T\left(v\right) - \vartheta_0\right) \Longrightarrow \zeta_v \sim \mathcal{N}\left(0, v^{-1}\mathrm{I}\left(\vartheta_0\right)^{-1}\right),$$

This estimator is used for adaptive Kalman filter.

2 Hidden AR

2.1 Model of observations

Consider the model of partially observed time series

$$X_t = f\,Y_{t-1} + \sigma\,w_t, \qquad X_0 \sim \mathcal{N}\left(0, d_x^2\right), \qquad t = 1, 2, \dots,$$
$$Y_t = a\,Y_{t-1} + b\,v_t, \qquad Y_0 \sim \mathcal{N}\left(0, d_y^2\right),$$

where $X^T = (X_0, X_1, \dots, X_T)$ are observations and auto regressive (AR) process $Y_t, t \geq 0$ is a hidden process. Here $v_t, w_t, t \geq 1$ are independent standard Gaussian random variables. The system is defined by the parameters $a, b, f, \sigma^2, d_x^2, d_y^2$. We study this model under condition

$$\mathscr{A}_0 \;: \qquad a^2 \in [0,1), \qquad b^2 > 0, \qquad f^2 > 0, \qquad \sigma^2 > 0.$$

It will be convenient for instant to denote $\vartheta = \left(a, b, f, \sigma^2\right)$.

The conditional expectation $m\left(\vartheta, t\right) = \mathbf{E}_\vartheta\left(Y_t | X_0, X_1, \ldots, X_t\right)$ according to the equations of Kalman filter (Kalman 1960) satisfies the equation

$$m\left(\vartheta, t\right) = a\, m\left(\vartheta, t-1\right) + \frac{af\gamma\left(\vartheta, t-1\right)}{\sigma^2 + f^2\gamma\left(\vartheta, t-1\right)}\left[X_t - fm\left(\vartheta, t-1\right)\right].$$

The initial value is $m\left(\vartheta, 0\right) = \mathbf{E}_\vartheta\left(Y_0 | X_0\right)$. The mean square error $\gamma\left(\vartheta, t\right) = \mathbf{E}_\vartheta\left(Y_t - m\left(\vartheta, t\right)\right)^2$ is described by the equation

$$\gamma\left(\vartheta, t\right) = a^2\gamma\left(\vartheta, t-1\right) + b^2 - \frac{a^2 f^2\gamma\left(\vartheta, t-1\right)^2}{\sigma^2 + f^2\gamma\left(\vartheta, t-1\right)}, \qquad t \geq 1$$

with the initial value $\gamma\left(\vartheta, 0\right) = \mathbf{E}_\vartheta\left(Y_0 - m\left(\vartheta, 0\right)\right)^2$.

If some of the mentioned parameters are unknown, then, of course, we can not use these equations for calculation of $m\left(\vartheta, t\right), t \geq 1$.

The proposed program is consists in several steps.

1. First on some learning interval $\left[0, \tau_T\right]$ of negligible length $(\tau_T/T \to 0)$ we construct a preliminary estimator $\vartheta^*_{\tau_T}$.

2. Then this estimator is used for defining the One-step MLE-process $\vartheta^\star_T = \left(\vartheta^\star_{t,T}, t = \tau_T + 1, \ldots, T\right)$

3. and finally the approximation $m^\star_T = \left(m^\star_{t,T}, t = \tau_T + 1, \ldots, T\right)$ is obtained by substituting $\vartheta^\star_T$ in the equations for $m\left(\vartheta, t\right)$.

4. The last step is to evaluate the error $m^\star_{t,T} - m\left(\vartheta, t\right)$.

Remark that the function $\gamma(\vartheta, t)$ converges to the value

$$\gamma_*(\vartheta) = \frac{f^2 b^2 - \sigma^2 \left(1 - a^2\right)}{2 f^2} + \frac{1}{2} \left[\left(\frac{\sigma^2 \left(1 - a^2\right)}{f^2} - b^2 \right)^2 + \frac{4 b^2 \sigma^2}{f^2} \right]^{1/2}$$

as $t \to \infty$. Therefore $m(\vartheta, t)$ satisfies

$$m_t(\vartheta) = a\, m_{t-1}(\vartheta) + \frac{a f \gamma_*(\vartheta)}{\sigma^2 + f^2 \gamma_*(\vartheta)} \left[X_t - f m_{t-1}(\vartheta)\right], \quad t \geq 1.$$

Note as well that if we denote ϑ_0 the true value, then

$$\zeta_t(\vartheta_0) = X_t - f_0 m_{t-1}(\vartheta_0) \left(\sigma_0^2 + f_0^2 \gamma_*(\vartheta_0)\right)^{-1/2}, \qquad t \geq 1$$

are i.i.d. standard Gaussian random variables. This means that the equation of observations is

$$X_t = f_0 m_{t-1}(\vartheta_0) + \sqrt{\sigma_0^2 + f_0^2 \gamma_*(\vartheta_0)}\, \zeta_t(\vartheta_0), \qquad t \geq 1.$$

Using this representation we can rewrite the equation for filter

$$m_t(\vartheta) = a\,m_{t-1}(\vartheta) + \frac{af\gamma_*(\vartheta)}{\sigma^2 + f^2\gamma_*(\vartheta)}\left[f_0 m_{t-1}(\vartheta_0) - f m_{t-1}(\vartheta)\right]$$

$$+ \frac{af\gamma_*(\vartheta)\sqrt{\sigma_0^2 + f_0^2\gamma_*(\vartheta_0)}}{\sigma^2 + f^2\gamma_*(\vartheta)}\zeta_t(\vartheta_0), \quad t \geq 1.$$

The likelihood function is

$$L\left(\vartheta, X^T\right) = \left(\frac{1}{2\pi P(\vartheta)}\right)^{T/2} \exp\left(-\frac{1}{2}\sum_{t=1}^{T}\frac{\left(X_t - f m_{t-1}(\vartheta)\right)^2}{P(\vartheta)}\right).$$

Here $P(\vartheta) = \sigma^2 + f^2\gamma_*(\vartheta)$ and Θ is an open, bounded, convex set of the possible values of the parameter ϑ.

2.2 Method of moments estimators

Introduce three statistics

$$S_{1,T}\left(X^T\right) = \frac{1}{T}\sum_{t=1}^{T}\left(X_t - X_{t-1}\right)^2,$$

$$S_{2,T}\left(X^T\right) = \frac{1}{T}\sum_{t=2}^{T}\left(X_t - X_{t-1}\right)\left(X_{t-1} - X_{t-2}\right),$$

$$S_{3,T}\left(X^T\right) = \frac{1}{T}\sum_{t=3}^{T}\left(X_t - X_{t-1}\right)\left(X_{t-2} - X_{t-3}\right)$$

and study their asymptotic ($T \to \infty$) behavior. We suppose that the condition $\mathscr{A}_0$ is fulfilled and the true value is denoted as ϑ_0.

Lemma 1. *We have the limits*

$$S_{1,T}\left(X^T\right) \quad \longrightarrow \quad \Phi_1\left(\vartheta_0\right)=\frac{2f^2b^2}{1+a}+2\sigma^2,$$

$$S_{2,T}\left(X^T\right) \quad \longrightarrow \quad \Phi_2\left(\vartheta_0\right)=\frac{f^2b^2\left(a-1\right)}{1+a}-\sigma^2,$$

$$S_{3,T}\left(X^T\right) \quad \longrightarrow \quad \Phi_3\left(\vartheta_0\right)=\frac{f^2b^2a\left(a-1\right)}{\left(1+a\right)},$$

and there exist constants $C_1>0$, $C_2>0$, $C_3>0$ *such that*

$$\sup_{\vartheta_0\in\mathbb{K}}\mathbf{E}_{\vartheta_0}\left|S_{1,T}\left(X^T\right)-\Phi_1\left(\vartheta_0\right)\right|^2\leq\frac{C_1}{T},$$

$$\sup_{\vartheta_0\in\mathbb{K}}\mathbf{E}_{\vartheta_0}\left|S_{2,T}\left(X^T\right)-\Phi_2\left(\vartheta_0\right)\right|^2\leq\frac{C_2}{T},$$

$$\sup_{\vartheta_0\in\mathbb{K}}\mathbf{E}_{\vartheta_0}\left|S_{3,T}\left(X^T\right)-\Phi_3\left(\vartheta_0\right)\right|^2\leq\frac{C_3}{T}.$$

Remark 1. We have as well the asymptotic normality

$$\sqrt{T}\left(S_{1,T}\left(X^T\right)-\Phi_1\left(\vartheta_0\right)\right)\Longrightarrow\mathcal{N}\left(0,D_1\left(\vartheta_0\right)^2\right),$$

$$\sqrt{T}\left(S_{2,T}\left(X^T\right)-\Phi_2\left(\vartheta_0\right)\right)\Longrightarrow\mathcal{N}\left(0,D_2\left(\vartheta_0\right)^2\right),$$

$$\sqrt{T}\left(S_{3,T}\left(X^T\right)-\Phi_3\left(\vartheta_0\right)\right)\Longrightarrow\mathcal{N}\left(0,D_3\left(\vartheta_0\right)^2\right)$$

All parameters of the model can be estimated with the help of the introduced statistics $S_{1,T}\left(X^T\right), S_{2,T}\left(X^T\right), S_{3,T}\left(X^T\right)$.

2.2.1 Estimation of the one-dim. parameters.

The MME of b has the properties

$$b_T^* = \left(\frac{\left(S_{1,T}\left(X^T\right) - 2\sigma^2\right)(1+a)}{2f^2} \right)^{1/2} \longrightarrow b_0$$

MME a_T^* is

$$a_T^* = \frac{2f^2b^2}{S_{1,T}\left(X^T\right) - 2\sigma^2} - 1 \longrightarrow a_0.$$

The MME

$$\sigma_T^{2*} = \frac{1}{2} S_{1,T}\left(X^T\right) - \frac{f^2b^2}{1+a} \longrightarrow \sigma_0^2.$$

2.2.2 Estimation of the parameter $\vartheta = (a, f)$.

The MME $\vartheta_T^* = (a_T^*, f_T^*)$ is solution of the system of equations

$$S_{1,T}\left(X^T\right) = \Phi_1\left(\vartheta_T^*\right), \qquad S_{2,T}\left(X^T\right) = \Phi_2\left(\vartheta_T^*\right)$$

and has the following form

$$a_T^* = \frac{S_{1,T}\left(X^T\right) + S_{2,T}\left(X^T\right) - \sigma^2}{S_{1,T}\left(X^T\right) - 2\sigma^2},$$

$$f_T^* = \left(\frac{S_{1,T}\left(X^T\right)\left(1 + a_T^*\right) - 2\sigma^2}{2b^2}\right)^{1/2}.$$

It was shown that

$$\left(a_T^*, f_T^*\right) \longrightarrow \left(a_0, f_0\right), \qquad \mathbf{E}_{\vartheta_0}\left\|\vartheta_T^* - \vartheta_0\right\|^2 \leq \frac{C}{T}.$$

2.2.3 Estimation of the parameter $\vartheta = \left(a, f, \sigma^2\right)$.

In this case we have three equations

$$S_{1,T}\left(X^T\right) = \frac{2f^2b^2}{1+a} + 2\sigma^2, \quad S_{2,T}\left(X^T\right) = \frac{f^2b^2\left(a-1\right)}{1+a} - \sigma^2,$$
$$S_{3,T}\left(X^T\right) = \frac{f^2b^2a\left(a-1\right)}{\left(1+a\right)}.$$

The MME $\vartheta_T^* = \left(a_T^*, f_T^*, \sigma_T^{2*}\right)$ is the following solution of this system:

$$a_T^* = \frac{2S_{3,T}\left(X^T\right)}{S_{1,T}\left(X^T\right) + 2S_{2,T}\left(X^T\right)} + 1,$$
$$f_T^* = \frac{S_{3,T}\left(X^T\right)\left(1+a_T^*\right)}{b^2a_T^*\left(a_T^*-1\right)},$$
$$\sigma_T^{2*} = \frac{1}{2}S_{1,T}\left(X^T\right) - \frac{\left(f_T^*\right)^2b^2}{1+a_T^*}.$$

2.3 Fisher informations

We have the same model of observations

$$X_t = fY_{t-1} + \sigma\, w_t, \qquad X_0, \qquad t \geq 1,$$
$$Y_t = aY_{t-1} + b\, v_t, \qquad Y_0, \qquad t \geq 1,$$

where $w_t, v_t, t \geq 1$ are independent standard Gaussian r.v.'s and f, σ^2, a, b are parameters of the model. As before, we suppose that some of these parameters are unknown and we have to estimate the unknown parameters $\vartheta \in \Theta$ by the observations $X^T = (X_0, X_1, \ldots, X_T)$. The MMEs studied above are consistent, but not asymptotically efficient. That is why we propose below the construction of One-step MLE-process, which allow us to solve two problems: first we obtain asymptotically efficient estimators of these parameters and the second - we describe the approximation of the conditional expectation $m\,(\vartheta, t)\,, t \geq 1$.

2.3.1 Unknown parameter b

Introduce the notation: $\Gamma\left(\vartheta\right)=f^2\gamma_*\left(\vartheta\right), P\left(\vartheta\right)=\sigma^2+\Gamma\left(\vartheta\right)$ and

$$A\left(\vartheta\right)=\frac{a\sigma^2}{P\left(\vartheta\right)},\qquad B\left(\vartheta,\vartheta_0\right)=\frac{a\Gamma\left(\vartheta\right)\sqrt{P\left(\vartheta_0\right)}}{P\left(\vartheta\right)},$$

$$\dot B_b\left(\vartheta_0,\vartheta_0\right)=\left.\frac{\partial B\left(\vartheta,\vartheta_0\right)}{\partial\vartheta}\right|_{\vartheta=\vartheta_0}=\frac{a\sigma^2\dot\Gamma\left(\vartheta_0\right)}{P\left(\vartheta_0\right)^{3/2}}.$$

Note that $\inf_{\vartheta\in\Theta}\dot\Gamma\left(\vartheta_0\right)>0$. Another important estimate is

$$\sup_{\vartheta\in\Theta}\left|A\left(\vartheta\right)\right|<1.$$

The Fisher information is

$$\mathrm{I}_b\left(\vartheta_0\right)=\frac{\dot P_b\left(\vartheta_0\right)^2\left[P\left(\vartheta_0\right)^2+a^2\sigma^4\right]}{2P\left(\vartheta_0\right)^2\left[P\left(\vartheta_0\right)^2-a^2\sigma^4\right]}.$$

2.3.2 Unknown parameter a

The Fisher information is

$$\mathrm{I}_a\left(\vartheta_0\right)=\frac{2Q\left(\vartheta_0\right)+\dot P\left(\vartheta_0\right)^2}{2P\left(\vartheta_0\right)},$$

where

$$Q\left(\vartheta_0\right)=$$

$$=\left(1-A\left(\vartheta_0\right)^2\right)^{-1}\left[\frac{\vartheta_0^2\Gamma\left(\vartheta_0\right)^2}{P\left(\vartheta_0\right)^2\left(1-\vartheta_0^2\right)}+\frac{\left[\Gamma\left(\vartheta_0\right)P\left(\vartheta_0\right)+\sigma^2\dot P\left(\vartheta_0\right)\right]^2}{P\left(\vartheta_0\right)^3}\right.$$

$$\left.+\frac{2A\left(\vartheta_0\right)\vartheta_0\Gamma\left(\vartheta_0\right)^2}{P\left(\vartheta_0\right)^2\left(1-\vartheta_0A\left(\vartheta_0\right)\right)}\left(\frac{\vartheta_0}{\left(1-A\left(\vartheta_0\right)^2\right)}+P\left(\vartheta_0\right)+\frac{\vartheta_0\sigma^2\dot P\left(\vartheta_0\right)}{\Gamma\left(\vartheta_0\right)}\right)\right]$$

2.4 One-step MLE-process

We consider the construction of One-step MLE-process in the case of unknown parameter $\vartheta = b \in \Theta = (\alpha_b, \beta_b)$,$\alpha_b > 0$. Let us fix a learning interval of observations $\left[X_0, X_1, \ldots, X_{\tau_T}\right]$, where $\tau_T = \left[T^\delta\right], \delta \in \left(\frac{1}{2}, 1\right)$ and $[A]$ here is the integer part of A. As preliminary estimator we take the MME $\vartheta^*_{\tau_T} = b^*_{\tau_T}$ defined by

$$\vartheta^*_{\tau_T} = 2^{-1/2} f^{-1} \left(\left[S_{1,\tau_T}\left(X^{\tau_T}\right) - 2\sigma^2\right](1+a)\right)^{1/2}.$$

One-step MLE-process is

$$\vartheta^\star_{t,T} = \vartheta^*_{\tau_T} + \frac{\mathrm{I}_b(\vartheta^*_{\tau_T})^{-1}}{(t-\tau_T)} \sum_{s=\tau_T+1}^{t} \left[\frac{\left[X_s - fm_{s-1}(\vartheta^*_{\tau_T})\right]}{P(\vartheta^*_{\tau_T})} f\dot{m}_{s-1}(\vartheta^*_{\tau_T}) \right.$$

$$\left. + \left(\left[X_s - fm_{s-1}(\vartheta^*_{\tau_T})\right]^2 - P(\vartheta^*_{\tau_T}) \right) \frac{\dot{P}(\vartheta^*_{\tau_T})}{2P(\vartheta^*_{\tau_T})^2} \right], \quad t \in \left[\tau_T + 2, T\right].$$

Theorem 1. *If* $t = [vT]\,, v \in (0,1]$ *and* $T \to \infty$, *then the following convergences*

$$\sqrt{t}\left(\vartheta^\star_{t,T} - \vartheta_0\right) \Longrightarrow \mathcal{N}\left(0, \mathrm{I}_b(\vartheta_0)^{-1}\right), \quad t\mathbf{E}_{\vartheta_0}\left(\vartheta^\star_{t,T} - \vartheta_0\right)^2 \longrightarrow \frac{1}{\mathrm{I}_b(\vartheta_0)}.$$

hold uniformly on compacts $\mathbb{K} \subset \Theta$.

Then the One-step MLE-process is given by the equality

$$\vartheta^{\star}_{t,T} = \theta^{*}_{\tau_T} + \frac{\mathbf{I}(\theta^{*}_{\tau_T})^{-1}}{t-\tau_T} \sum_{s=\tau_T+1}^{t} \left[\frac{\left[X_s - M(\vartheta^{*}_{\tau_T}, s-1)\right]}{P(\vartheta^{*}_{\tau_T})} \dot{\mathrm{M}}(\vartheta^{*}_{\tau_T}, s-1) \right.$$

$$\left. + \left(\left[X_s - M(\vartheta^{*}_{\tau_T}, s-1)\right]^2 - P(\vartheta^{*}_{\tau_T}) \right) \frac{\dot{\mathrm{P}}(\vartheta^{*}_{\tau_T})}{2P(\vartheta^{*}_{\tau_T})^2} \right], \; t \in \left[\tau_T + 2, T\right].$$

It can be shown that for any $v \in (0, T]$, if we put $t = vT$ then the normalized difference is asymptotically normal:

$$\sqrt{t}\left(\vartheta^{\star}_{t,T} - \vartheta_0\right) \Longrightarrow \xi\left(\vartheta_0\right) \sim \mathcal{N}\left(0, \mathbf{I}(\theta_0)^{-1}\right),$$

$$t\mathbf{E}_{\vartheta_0}\left\|\vartheta^{\star}_{t,T} - \vartheta_0\right\|^2 \longrightarrow \mathbf{E}_{\vartheta_0}\left\|\xi\left(\vartheta_0\right)\right\|^2.$$

2.5 MLE and BE

Consider the model of observations

$$X_t = f\,Y_{t-1} + \sigma\, w_t, \qquad X_0 \sim \mathcal{N}\left(0, d_x^2\right), \qquad t = 1, 2, \ldots,$$

$$Y_t = a\,Y_{t-1} + b\, v_t, \qquad Y_0 \sim \mathcal{N}\left(0, d_y^2\right),$$

where the unknown parameter is $\vartheta = b \in \Theta = (\alpha_b, \beta_b)\,, \alpha_b > 0$. Below we study the MLE $\hat{\vartheta}_T$ and BE $\tilde{\vartheta}_T$ defined by the usual relations

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

Here $p\,(\vartheta)\,, \vartheta \in \Theta$ is continuous, positive on Θ density a priory.

Recall the notation $P(\vartheta) = \sigma^2 + f^2\gamma_*(\vartheta)$,

$$A(\vartheta) = \frac{a\sigma^2}{P(\vartheta)^2}, \qquad B(\vartheta,\vartheta_0) = \frac{af\gamma_*(\vartheta)\sqrt{P(\vartheta_0)}}{P(\vartheta)},$$

$$\dot B(\vartheta_0,\vartheta_0) = \left.\frac{\partial B(\vartheta,\vartheta_0)}{\partial\vartheta}\right|_{\vartheta=\vartheta_0} = \frac{af\sigma^2\dot\gamma_*(\vartheta_0)}{P(\vartheta_0)^{3/2}}.$$

$$\mathrm{I}_b(\vartheta_0) = \frac{f^2\dot B(\vartheta_0,\vartheta_0)^2}{\left[1 - A(\vartheta_0)^2\right]P(\vartheta_0)} + \frac{\dot P(\vartheta_0)^2}{2P(\vartheta_0)^2}$$

$$= \frac{f^2\dot\gamma_*(\vartheta)^2}{P(\vartheta_0)}\left[\frac{a^2\sigma^4}{P(\vartheta_0)^2 - a^2\sigma^4} + \frac{1}{2P(\vartheta_0)}\right].$$

Below it is shown that the family of measures is LAN. Therefore

We have Hajek-Le Cam's lower bound: for any $\bar\vartheta_T$ and any $\vartheta_0 \in \Theta$

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

This bound allows us to define the asymptotically efficient estimator ϑ_T° as estimator for which the equality holds for all $\vartheta_0 \in \Theta$.

Theorem 2. *The MLE and BE are consistent, asymptotically normal,*

$$\sqrt{T}\left(\hat\vartheta_T - \vartheta_0\right) \Longrightarrow \mathcal{N}\left(0, \mathrm{I}_b\left(\vartheta_0\right)^{-1}\right),$$

$$\sqrt{T}\left(\tilde\vartheta_T - \vartheta_0\right) \Longrightarrow \mathcal{N}\left(0, \mathrm{I}_b\left(\vartheta_0\right)^{-1}\right),$$

the polynomial moments converge and the both estimators are asymptotically efficient.

2.6 Adaptive filter

We are given the partially observed system

$$
\begin{aligned}
X_t &= f\,Y_{t-1} + \sigma\, w_t, \qquad X_0, \qquad t = 1, 2, \ldots,\\
Y_t &= a\,Y_{t-1} + b\, v_t, \qquad Y_0.
\end{aligned}
$$

Recall that if all parameters of the model $\vartheta = \left(f, \sigma^2, a, b\right)$ are known, then the stationary version $m_t(\vartheta)$ of the conditional expectation $m(\vartheta, t) = \mathbf{E}_\vartheta \left(Y_t | X_s, s \leq t\right)$ satisfies the equation

$$
\begin{aligned}
m_t(\vartheta) &= a m_{t-1}(\vartheta) + \frac{a f \gamma_*(\vartheta)}{\sigma^2 + f^2 \gamma_*(\vartheta)} \left[X_t - f m_{t-1}(\vartheta)\right], \quad m_0(\vartheta),\\
\gamma_*(\vartheta) &= \frac{1}{2f^2}\left[f^2 b^2 - \sigma^2\left(1 - a^2\right)\right]\\
&\qquad + \frac{1}{2f^2 \sqrt{\left(\sigma^2\left(1-a^2\right) - b^2 f^2\right)^2 + 4 f^2 b^2 \sigma^2}}.
\end{aligned}
$$

2.6.1 Unknown parameter b

Suppose that the values $f > 0$, $a^2 \in [0,1)$ and $\sigma^2 > 0$ are known and the only unknown parameter is $\vartheta = b \in \Theta = (\alpha_b, \beta_b)$, $\alpha_b > 0$. The $\vartheta^*_{\tau_T} = b^*_{\tau_T}$, $\tau_T = \left[T^\delta\right], \delta \in (\frac{1}{2}, 1)$.

$$\vartheta^*_{\tau_T} = 2^{-1/2} f^{-1} \left[\left(\frac{1}{\tau_T}\sum_{s=1}^{\tau_T}\left[X_s - X_{s-1}\right]^2 + \sigma^2\right)(1+a)\right]^{1/2}.$$

Then we define the One-step MLE-process

$$\vartheta^\star_{t,T} = \vartheta^*_{\tau_T} + \frac{\mathrm{I}_b(\vartheta^*_{\tau_T})^{-1}}{(t-\tau_T)} \sum_{s=\tau_T+1}^{t} \left[\frac{\left[X_s - f m_{s-1}(\vartheta^*_{\tau_T})\right]}{P(\vartheta^*_{\tau_T})} f\dot m_{s-1}(\vartheta^*_{\tau_T})\right.$$
$$\left. + \left(\left[X_s - f m_{s-1}(\vartheta^*_{\tau_T})\right]^2 - P(\vartheta^*_{\tau_T})\right)\frac{\dot P(\vartheta^*_{\tau_T})}{2P(\vartheta^*_{\tau_T})^2}\right],\ t \in \left[\tau_T + 1, T\right].$$

For $s > \tau_T + 1$ we can write

$$m_{s-1}(\vartheta^*_{\tau_T}) = P(\vartheta^*_{\tau_T})^{-1}\left[a\sigma^2 m_{s-2}(\vartheta^*_{\tau_T}) + af\gamma_*(\vartheta^*_{\tau_T})X_{s-1}\right],$$

$$\dot m_{s-1}(\vartheta^*_{\tau_T}) = \frac{a\sigma^2\left[P(\vartheta^*_{\tau_T})\dot m_{s-2}(\vartheta^*_{\tau_T}, s-2) - f^2\dot\gamma_*(\vartheta^*_{\tau_T})\, m_{s-2}(\vartheta^*_{\tau_T})\right]}{P(\vartheta^*_{\tau_T})^2} + \frac{af\sigma^2\dot\gamma_*\left(\vartheta\right)}{P\left(\vartheta\right)^2}\, X_{s-1}.$$

Here $P\left(\vartheta\right) = \sigma^2 + f^2\gamma_*\left(\vartheta\right)$ and the Fisher information is

$$\mathrm{I}_b(\vartheta) = \frac{\dot P_b\left(\vartheta_0\right)^2\left[P\left(\vartheta_0\right)^2 + a^2\sigma^4\right]}{2P\left(\vartheta_0\right)^2\left[P\left(\vartheta_0\right)^2 - a^2\sigma^4\right]},$$

where $A\left(\vartheta\right) = P\left(\vartheta\right)^{-1} a\sigma^2$.

The adaptive Kalman filter we introduce with the help of the process $m^\star_{t,T}$ defined as follows

$$\hat m_{t,T} = P(\vartheta^\star_{t-1,T})^{-1}\left[a\sigma^2 m^\star_{t-1,T} + af\gamma_*(\vartheta^\star_{t-1,T})X_t\right], \quad t\in\left[\tau_T+1,T\right]$$

We have to compare $m^\star_{t,T}$ with $m\left(\vartheta_0,t\right)$ for large T.

Below $\dot B\left(\vartheta_0,\vartheta_0\right) = -f^{-1}\sqrt{P\left(\vartheta_0\right)}\dot A(\vartheta_0) = P\left(\vartheta_0\right)^{-3/2}af\sigma^2\dot\gamma_*\left(\vartheta_0\right)$ and $\eta_{t,T}=\sqrt{t}\left[\vartheta^\star_{t,T}-\vartheta_0\right]$.

Theorem 3. *Let $t=[vT]\,, v\in(0,1]$, $k=t-\tau_T+2$ and $T\to\infty$. Then the following relations hold*

$$\sqrt{t}\left[m^\star_{t,T}-m_t\left(\vartheta_0\right)\right]$$

$$= \dot B\left(\vartheta_0,\vartheta_0\right)\sum_{m=0}^{k} A\left(\vartheta_0\right)^m \eta_{t-m-1,T}\,\zeta_{t-m}\left(\vartheta_0\right)+o\left(1\right),$$

$$t\mathbf{E}_{\vartheta_0}\left[m^\star_{t,T}-m_t\left(\vartheta_0\right)\right]^2 \longrightarrow \frac{\dot B\left(\vartheta_0,\vartheta_0\right)^2}{\mathrm{I}_b\left(\vartheta_0\right)\left(1-A\left(\vartheta_0\right)^2\right)}.$$

Remark 2. The adaptive Kalman filter is given by the relations, where the only One-step MLE-process is in non recurrent form. Let us write this estimator in recurrent form too

$$\vartheta^{\star}_{t,T} = \frac{\vartheta^{*}_{\tau_T}}{t-\tau_T} + \left(1 - \frac{1}{t-\tau_T}\right) \vartheta^{\star}_{t-1,T}$$

$$+ \frac{\left[X_t - f m_{t-1}(\vartheta^{*}_{\tau_T})\right] f \dot m_{t-1}(\vartheta^{*}_{\tau_T})}{\mathrm{I}_b(\vartheta^{*}_{\tau_T})\left(t-\tau_T\right) P(\vartheta^{*}_{\tau_T})}$$

$$+ \frac{\left(\left[X_t - f m_{t-1}(\vartheta^{*}_{\tau_T})\right]^2 - P(\vartheta^{*}_{\tau_T})\right) \dot P(\vartheta^{*}_{\tau_T})}{2\mathrm{I}_b(\vartheta^{*}_{\tau_T})\left(t-\tau_T\right) P(\vartheta^{*}_{\tau_T})^2}, \quad t \in \left[\tau_T + 1, T\right].$$

Now the adaptive Kalman filter is in recurrent form.

Remark 3. The cases $\vartheta = f$, $\vartheta = a$ and $\vartheta = \sigma^2$ can be studied similarly. If we consider the two-dimensional cases, say, $\vartheta = (f, a)$, then the corresponding adaptive Kalman filter can be written as well. The preliminary MME estimator $\vartheta^*_{\tau_T} = (f^*_T, a^*_T)^\top$ was defined above, for the Fisher information matrix see above too, the equations for $\dot{m}_f\left(\vartheta_T, t\right)$ and $\dot{m}_a\left(\vartheta_T, t\right)$ can be easily written. The equation for $m^\star_{t,T}$ has exactly the same form as the given above.

2.7 Asymptotic efficiency

We have the same system

$$X_t = f\,Y_{t-1} + \sigma\,w_t, \qquad X_0, \qquad t = 1, 2, \ldots,$$
$$Y_t = a\,Y_{t-1} + b\,v_t, \qquad Y_0.$$

where the parameters f, a, b, σ^2 satisfy the condition $\mathscr{A}_0$.

The adaptive filter is given above and we would like to know if the error of approximation $\mathbf{E}_\vartheta \left|m^\star_{t,T} - m\,(\vartheta, t)\right|^2$ is asymptotically minimal? As before we propose a lower minimax bound on the risks of all estimators $\bar m_t$ supposing that their calculation is based on the observations up to time t, i.e., these estimators depend on $X^t = (X_s, 0 \leq s \leq t)$.

Denote the limit (below $\eta^\star_{vT,T}\left(\vartheta_0\right)=\sqrt{vT}\left(\tilde\vartheta_{vT,T}-\vartheta_0\right)$)

$$S\left(\vartheta_0\right)^2=\lim_{T\rightarrow\infty}\mathbf{E}_{\vartheta_0}\left[\dot m\left(\vartheta_0,vT\right)^2\eta^\star_{vT,T}\left(\vartheta_0\right)^2\right].$$

The value of $S\left(\vartheta_0\right)^2$ can be calculated with the help of the representations for $m\left(\vartheta\right)$ and $\dot m\left(\vartheta\right)$, but the corresponding expression will be too cumbersome and we omit this expression.

Theorem 4. *Let the conditions of Theorem 1 be fulfilled. Then we have the following lower minimax bound: for any estimator $\bar m_{t,T}$ of $m\left(\vartheta,t\right)$ (below $t=vT$)*

$$\lim_{\nu\rightarrow 0}\varliminf_{T\rightarrow\infty}\sup_{\left|\vartheta-\vartheta_0\right|\leq\nu}t\mathbf{E}_\vartheta\left|\bar m_{t,T}-m\left(\vartheta,t\right)\right|^2\geq S\left(\vartheta_0\right)^2.$$

We call the estimator-process $m^{\circ}_{t,T}, \tau_T < t \leq T$ *asymptotically efficient* if for all $\vartheta_0 \in \Theta$, $t = vT$, any $v \in [\varepsilon_0, 1]$

$$\lim_{\nu \to 0} \lim_{T \to \infty} \sup_{|\vartheta - \vartheta_0| \leq \nu} t\mathbf{E}_{\vartheta} \left| m^{\circ}_{t,T} - m(\vartheta, t) \right|^2 = S(\vartheta_0)^2 .$$

Here $\varepsilon_0 \in (0, 1)$.

Theorem 5. *The estimator $m^{\star}_t, \tau_T < t \leq T$ is asymptotically efficient.*