МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО
ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Международная лаборатория статистики случайных
процессов и количественного финансового анализа

Международная научная конференция «Робастная статистика и финансовая математика – 2022»

(04–05 июля 2022 г.)

Сборник статей

Под редакцией
д-ра физ.-мат. наук, профессора С.М. Пергаменщикова,
канд. физ.-мат. наук, доцента Е.А. Пчелинцева

Томск
Издательский Дом Томского государственного университета
2022

Statistics of hidden Markov processes (continuous time) [1]

Kutoyants Yu.A.

Le Mans University, Le Mans, France,
Tomsk State University, Tomsk, Russia,
e-mail: Yury.Kutoyants@univ-lemans.fr

Abstract

We present a survey of several recent results on parameter estimation for partially observed linear systems and the construction of adaptive Kalman-Bucy filtration equations. It is supposed that the observation equation contains small noise of level ε and the properties of estimators are described in the asymptotics $\varepsilon \to 0$. The adaptive filter is constructed in several steps. First we propose a nonparametric estimator of the quadratic variation of the derivative of the observations. Then we use this estimator for construction of One-step MLE-process. Finally, this estimator-process is substituted in the filtration equations.

Keywords: Kalman-Bucy filter; Volatility estimation; Adaptive filtration; One-step MLE-process.

MLE and BE.

Consider a partially observed linear system

$$\mathrm{d}X_t = f(\vartheta,t)\,Y_t \mathrm{d}t + \varepsilon\sigma(t)\,\mathrm{d}W_t, \quad X_0 = 0, \quad 0 \le t \le T \tag{1}$$

$$\mathrm{d}Y_t = a(\vartheta,t)\,Y_t \mathrm{d}t + b(\vartheta,t)\,\mathrm{d}V_t, \quad Y_0 = y_0, \tag{2}$$

where $f(\cdot), \sigma(\cdot), a(\cdot)$ and $b(\cdot)$ are known, smooth functions, while W_t, and V_t are two independent Wiener processes. We have to estimate $\vartheta \in \Theta = (\alpha,\beta)$ from continuous time observations X^T, given that the process $(Y_t, 0 \le t \le T)$ is unobservable (hidden).

The likelihood ratio function is

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

Here $m(\vartheta,t) = \mathbf{E}_\vartheta\left(Y_t | X_s, 0 \le s \le t\right)$ is solution of the equations:

Kalman–Bucy filter.

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

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

[1] The work was supported by the RSF, project no 20-61-47043.

where $\gamma(\vartheta,t)=\mathbf{E}_\vartheta\left(m(\vartheta,t)-Y_t\right)^2$.

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

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

The model under study is of special interest when $\varepsilon\to 0$ because,

$$\lim_{\varepsilon\to 0}\{f(\vartheta,t)\,\dot m(\vartheta,t)+\dot f(\vartheta,t)\,m(\vartheta,t)\}=0,$$

$$\varepsilon^2\mathrm{I}_\varepsilon(\vartheta)=\int_0^T\sigma(t)^{-2}\left[\frac{\partial}{\partial\vartheta}\{f(\vartheta,t)\,m(\vartheta,t)\}\right]^2\mathrm{d}t\to 0.$$

In addition we have the convergence in distribution

$$\varepsilon^{-1/2}\{f(\vartheta,t)\,\dot m(\vartheta,t)+\dot f(\vartheta,t)\,m(\vartheta,t)\}\Longrightarrow n(\vartheta,t)\;\xi_t,$$

where $n(\vartheta,t)$ is a det. f. and ξ_t are i.i.d. $\mathcal{N}(0,1/2)$. We expect

$$\varepsilon\,\mathrm{I}_\varepsilon(\vartheta)\Longrightarrow\int_0^T n(\vartheta,t)^2\,\xi_t^2\,\mathrm{d}t.$$

As it happens, however, this integral does not exist and we have instead, as $\varepsilon\to 0$,

$$\varepsilon\,\mathrm{I}_\varepsilon(\vartheta)\longrightarrow\mathrm{I}_0(\vartheta)=\frac12\int_0^T n(\vartheta,t)^2\;\mathrm{d}t.$$

Limit model. Suppose that $\varepsilon=0$, then we obtain the system

$$\mathrm{d}x_t=f(\vartheta,t)\,Y_t\mathrm{d}t,\qquad x_t'=f(\vartheta,t)\,Y_t,\quad x_0'=f(\vartheta,0)\,y_0,$$

$$\mathrm{d}Y_t=a(\vartheta,t)\,Y_t\mathrm{d}t+b(\vartheta,t)\,\mathrm{d}V_t,\quad Y_0=y_0$$

Question: *is it possible to estimate ϑ by observations $x_t',0\le t\le T$ without error?*

Recall that by Itô's formula, we have

$$\left(x_t'\right)^2=2\int_0^t x_s'\mathrm{d}x_s'+\int_0^t b(\vartheta,s)^2\,f(\vartheta,s)^2\,\mathrm{d}s$$

and the function

$$\Psi_t=\left(x_t'\right)^2-2\int_0^t x_s'\mathrm{d}x_s'=\int_0^t b(\vartheta,s)^2\,f(\vartheta,s)^2\,\mathrm{d}s\equiv K(\vartheta,t)$$

is deterministic. The "observed" function Ψ_t defines ϑ without error. For example, the estimator ϑ^* defined by the equation $K(\vartheta^*,t)=\Psi_t$ is without error, i.e., $\vartheta^*=\vartheta$.

Let us denote $S(\vartheta,t)=f(\vartheta,t)\,b(\vartheta,t)$, set

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

$$\mathrm{I}_0(\vartheta)=\int_0^T\frac{S(\vartheta,t)}{2\sigma(t)}\left[\frac{\partial}{\partial\vartheta}\ln S(\vartheta,t)\right]^2\mathrm{d}t$$

and introduce the conditions:

For any $\nu>0$,

$$\inf_{\vartheta_0\in\Theta}\inf_{|\vartheta-\vartheta_0|>\nu}G(\vartheta,\vartheta_0)>0\quad\text{and}\quad\inf_{\vartheta\in\Theta}\mathrm{I}_0(\vartheta)>0.$$

We have the lower minimax bound

$$\lim_{\delta\to 0}\varliminf_{\varepsilon\to 0}\sup_{|\vartheta-\vartheta_0|<\delta}\varepsilon^{-1}\mathbf{E}_\vartheta|\vartheta_\varepsilon^*-\vartheta|^2=\mathrm{I}_0\left(\vartheta\right)^{-1}.$$

Theorem 1. *The MLE $\hat\vartheta_\varepsilon$ and the BE $\tilde\vartheta_\varepsilon$ are consistent, and asymptotically normal, i.e.,*

$$\frac{\hat\vartheta_\varepsilon-\vartheta_0}{\sqrt{\varepsilon}}\Longrightarrow\zeta\sim\mathcal{N}\left(0,\mathrm{I}_0\left(\vartheta_0\right)^{-1}\right),\qquad\frac{\tilde\vartheta_\varepsilon-\vartheta_0}{\sqrt{\varepsilon}}\Longrightarrow\zeta.$$

Moreover the moments converge, i.e., for any $p>0$

$$\lim_{\varepsilon\to 0}\mathbf{E}_{\vartheta_0}\left|\frac{\hat\vartheta_\varepsilon-\vartheta_0}{\sqrt{\varepsilon}}\right|^p=\mathbf{E}_{\vartheta_0}\left|\zeta\right|^p,\qquad\lim_{\varepsilon\to 0}\mathbf{E}_{\vartheta_0}\left|\frac{\tilde\vartheta_\varepsilon-\vartheta_0}{\sqrt{\varepsilon}}\right|^p=\mathbf{E}_{\vartheta_0}\left|\zeta\right|^p$$

and both estimators are asymptotically efficient.

For the proof see [1].

Quadratic variation estimation. Let us consider the linear two-dimensional partially observed system

$$\mathrm{d}X_t=f\left(t\right)Y_t\mathrm{d}t+\varepsilon\sigma\left(t\right)\mathrm{d}W_t,\qquad X_0=0,\quad 0\leq t\leq T,\tag{3}$$

$$\mathrm{d}Y_t=a\left(t\right)Y_t\mathrm{d}t+b\left(t\right)\mathrm{d}V_t,\quad Y_0=0,\quad 0\leq t\leq T,\tag{4}$$

where the Wiener processes $V_t, 0\leq t\leq T$ and $W_t, 0\leq t\leq T$ are supposed to be independent. The solution $Y^T=(Y_t, 0\leq t\leq T)$ can not be observed directly and we have available the observations $X^T=(X_t, 0\leq t\leq T)$ only. Here $a\left(\cdot\right), b\left(\cdot\right), f\left(\cdot\right), \sigma\left(\cdot\right)\in\mathcal{C}^1$ are unknown functions and $\varepsilon\in(0,1]$ is a *small* parameter.

Our first goal is to construct a consistent $(\varepsilon\to 0)$ estimator $\hat\Psi_{\tau,\varepsilon}, 0\leq\tau\leq T$ of the function

$$\Psi_\tau=\int_0^\tau f\left(s\right)^2b\left(s\right)^2\mathrm{d}s,\qquad 0<\tau\leq T.\tag{5}$$

Then we show that this estimator $\hat\Psi_{\tau,\varepsilon}$ can be useful in the construction of the estimators of the parameters ϑ in the case of models (1), (2) with $f\left(t\right)=f\left(\vartheta,t\right)$ or $b\left(t\right)=b\left(\vartheta,t\right)$.

The construction of the estimators is based on the following properties of the model (3), (4).

$$\sup_{0\leq t\leq T}\left|X_t-x_t\right|\longrightarrow 0,\qquad x_t=\int_0^t f\left(s\right)Y_s\,\mathrm{d}s.$$

We have

$$\frac{\mathrm{d}x_t}{\mathrm{d}t}=f\left(t\right)Y_t,\qquad x_0=0.$$

If we formally calculate the derivative

$$\left.\frac{\partial X_t}{\partial t}\right|_{\varepsilon=0}=f_t\left(\vartheta\right)Y_t$$

and then calculating the quadratic variation $\left\langle f_t\left(\tau\right)Y_\tau\right\rangle=\Psi_\tau$ we obtain the desired function Ψ_τ. The construction of the estimator of Ψ_τ is a discrete time modification of these two steps.

Introduce the statistic

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

Here $t_i=i\varphi_\varepsilon, N_{\tau,\varepsilon}=\left[\frac{\tau}{\varphi_\varepsilon}\right]$, the rates $\varphi_\varepsilon\rightarrow 0,\delta_\varepsilon\rightarrow 0$ will be defined later.

Let us explain why this statistic can be a consistent estimator of Ψ_τ. We have

$$\frac{X_{t_{i+1}+\delta_\varepsilon}-X_{t_{i+1}}}{\delta_\varepsilon}=\frac{1}{\delta_\varepsilon}\int_{t_{i+1}}^{t_{i+1}+\delta_\varepsilon}f_s(\vartheta)\,Y_s\,\mathrm{d}s+\frac{\varepsilon}{\delta_\varepsilon}\int_{t_{i+1}}^{t_{i+1}+\delta_\varepsilon}\sigma_s\,\mathrm{d}W_s.$$

Hence if we take $\varepsilon\delta_\varepsilon^{-1/2}\rightarrow 0$, then

$$\frac{X_{t_{i+1}+\delta_\varepsilon}-X_{t_{i+1}}}{\delta_\varepsilon}=f_{t_{i+1}}(\vartheta)\,Y_{t_{i+1}}+o(1),$$

$$\frac{X_{t_i+\delta_\varepsilon}-X_{t_i}}{\delta_\varepsilon}=f_{t_i}(\vartheta)\,Y_{t_i}+o(1)$$

and

$$\left|f_{t_{i+1}}(\vartheta)-f_{t_i}(\vartheta)\right|\leq C\varphi_\varepsilon.$$

Further, formally we write

$$\begin{aligned}\Psi_\varepsilon(\tau)&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}\left(f_{t_{i+1}}(\vartheta)\,Y_{t_{i+1}}-f_{t_i}(\vartheta)\,Y_{t_i}\right)^2+o(1)\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_{t_i}(\vartheta)^2\left(Y_{t_{i+1}}-Y_{t_i}\right)^2+o(1)\\&=\sum_{i=0}^{N_{t,\varepsilon}-1}f_{t_i}(\vartheta)^2\left(\int_{t_i}^{t_{i+1}}a_s(\vartheta)\,Y_s\mathrm{d}s+\int_{t_i}^{t_{i+1}}b_s(\vartheta)\,\mathrm{d}V_s\right)^2+o(1)\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_{t_i}(\vartheta)^2\left(\int_{t_i}^{t_{i+1}}b_s(\vartheta)\,\mathrm{d}V_s\right)^2+o(1)\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_{t_i}(\vartheta)^2\,b_{t_i}(\vartheta)^2\;(t_{i+1}-t_i)+o(1)\\&\longrightarrow\int_0^\tau f_s(\vartheta)^2\,b_s(\vartheta)^2\,\mathrm{d}s=\Psi_\tau.\end{aligned}$$

We have the estimates

$$\begin{aligned}\mathbf{E}_\vartheta&\left|\Psi_\varepsilon(\tau)-\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_{t_i}(\vartheta)^2\left[Y_{t_{i+1}}-Y_{t_i}\right]^2\right|\\&\qquad\qquad\leq C\left[\sqrt{\frac{\delta_\varepsilon}{\varphi_\varepsilon}}+\varphi_\varepsilon+\frac{\varepsilon}{\sqrt{\delta_\varepsilon\varphi_\varepsilon}}\right],\end{aligned}$$

and

$$\left|\Psi_\tau - \sum_{i=0}^{N_{\tau,\varepsilon}-1} f_{t_i}(\vartheta)^2 b_{t_i}(\vartheta)^2 (t_{i+1}-t_i)\right| \leq C\varphi_\varepsilon.$$

If we put $\delta_\varepsilon = \varepsilon^q, \varphi_\varepsilon = \varepsilon^l$ then the equation

$$\sqrt{\frac{\delta_\varepsilon}{\varphi_\varepsilon}} = \varphi_\varepsilon = \frac{\varepsilon}{\sqrt{\delta_\varepsilon\varphi_\varepsilon}}$$

gives us the values $q = 1$ and $l = 1/3$, i.e.,

$$\delta_\varepsilon = \varepsilon \qquad \text{and} \qquad \varphi_\varepsilon = \varepsilon^{1/3}.$$

Hence $t_{i+1} = t_i + \varepsilon^{1/3}$,

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

and

$$\mathbf{E}_\vartheta \left|\Psi_\varepsilon(\tau) - \sum_{i=0}^{N_{\tau,\varepsilon}-1} f_{t_i}(\vartheta)^2 \left[Y_{t_{i+1}} - Y_{t_i}\right]^2\right| \leq C\,\varepsilon^{1/3}.$$

The regularity conditions:

$\mathcal{A}$. The functions $a(\cdot), b(\cdot) \in \mathcal{C}^{(1)}[0,\tau]$ and $f(\cdot), \sigma(\cdot) \in \mathcal{C}^{(2)}[0,\tau]$.

Theorem 2. *Let the condition $\mathcal{A}$ be fulfilled then for any $p > 0$ there exist a constant $C > 0$ such that*

$$\mathbf{E}_\vartheta \left|\hat\Psi_\varepsilon(\tau) - \Psi_\tau\right|^p \leq C\,\varepsilon^{p/3}.$$

Reark that it is possible to prove the asymptotic normality

$$\varepsilon^{-1/3}\left(\hat\Psi_\varepsilon(\tau) - \Psi_\tau\right) \Longrightarrow \mathcal{N}\left(0, D^2\right).$$

Parameter estimation.

Consider the partially observed linear system

$$\mathrm{d}Y_t = a(t)\,Y_t \mathrm{d}t + b(\vartheta,t)\,\mathrm{d}V_t, \qquad Y_0 = 0, \tag{6}$$

$$\mathrm{d}X_t = f(\vartheta,t)\,Y_t \mathrm{d}t + \varepsilon\sigma(t)\,\mathrm{d}W_t, \qquad X_0 = 0. \tag{7}$$

Substitution estimator $\bar\vartheta_\varepsilon$.

The functions $f(\cdot)$, $b(\cdot)$ are supposed to be known and the functions $a(\cdot)$, $\sigma(\cdot)$ are unknown. Define the function

$$\Psi_\tau(\vartheta) = \int_0^\tau f(\vartheta,t)^2 b(\vartheta,t)^2\,\mathrm{d}t, \qquad \vartheta \in (\alpha,\beta) = \Theta.$$

We have

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

Condition $\mathcal{B}$.

$\mathcal{B}_1$. The functions $\Psi_\tau(\vartheta)$ has two continuous derivatives $\dot\Psi_\tau(\vartheta)$, $\ddot\Psi_\tau(\vartheta)$ w.r.t. ϑ.

$\mathcal{B}_2$. *For a given τ we have*

$$\inf_{\vartheta\in\Theta}\left|\dot\Psi_\tau\left(\vartheta\right)\right|>0. \tag{8}$$

By condition (8) the function $\Psi_\tau\left(\cdot\right)$ is monotone. Without loss of generality we suppose that it is increasing.

Introduce the notation

$$\hat\Psi_{\tau,\varepsilon}=\sum_{i=0}^{N_{\tau,\varepsilon}-1}\left(\frac{X_{t_{i+1}+\varepsilon}-X_{t_{i+1}}}{\varepsilon}-\frac{X_{t_i+\varepsilon}-X_{t_i}}{\varepsilon}\right)^2,\quad t_{i+1}=t_i+\varepsilon^{1/3},$$

$$\psi_m=\inf_{\vartheta\in\Theta}\Psi_\tau\left(\vartheta\right),\ \psi_M=\sup_{\vartheta\in\Theta}\Psi_\tau\left(\vartheta\right),\ \psi_m=\Psi_\tau\left(\alpha\right),\ \psi_M=\Psi_\tau\left(\beta\right),$$

$$G\left(\psi\right)=\Psi_\tau^{-1}\left(\psi\right),\quad \psi_m<\psi<\psi_M,\qquad \alpha<G\left(\psi\right)<\beta,$$

$$\mathbb{B}_m=\left\{\omega:\quad \hat\Psi_{\tau,\varepsilon}\leq\psi_m\right\},\qquad \mathbb{B}_M=\left\{\omega:\quad \hat\Psi_{\tau,\varepsilon}\geq\psi_M\right\},$$

$$\mathbb{B}=\left\{\omega:\quad \psi_m<\hat\Psi_{\tau,\varepsilon}<\psi_M\right\},\qquad \eta_\varepsilon=G(\hat\Psi_{\tau,\varepsilon}),$$

The substitution estimator (SE) is introduced as follows

$$\check\vartheta_{\tau,\varepsilon}=\alpha1\!\!1_{\{\mathbb{B}_m\}}+\eta_\varepsilon1\!\!1_{\{\mathbb{B}\}}+\beta1\!\!1_{\{\mathbb{B}_M\}}.$$

Roughly speaking

$$\hat\Psi_{\tau,\varepsilon}=\Psi_\tau\left(\check\vartheta_{\tau,\varepsilon}\right).$$

It has the following properties.

Theorem 3. *Suppose that the conditions $\mathcal{A}$ and $\mathcal{B}$ are fulfilled. Then for any $p>0$ there exists a constant $C=C\left(p\right)>0$ such that*

$$\sup_{\vartheta\in\Theta}\varepsilon^{-p/3}\mathbf{E}_\vartheta\left|\check\vartheta_{\tau,\varepsilon}-\vartheta\right|^p\leq C. \tag{9}$$

Example 1. Suppose that we have the model of observations (6),(7), where $f\left(\vartheta,t\right)=\vartheta f\left(t\right),\vartheta\in\left(\alpha,\beta\right),\alpha>0$, $b\left(\vartheta,t\right)=b\left(t\right)$ and all corresponding conditions are fulfilled. Then

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

and SE

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

This estimator is consistent and has the rate of convergence $\varepsilon^{1/3}$.

Example 2. Consider this model with $b\left(\vartheta,t\right)=\sqrt{h\left(t\right)+\vartheta g\left(t\right)}$ and $f\left(\vartheta,t\right)=f\left(t\right)$. Suppose that the functions $h\left(\cdot\right)$ and $g\left(\cdot\right)$ are positive and $\alpha>0$. Then

$$\Psi_\tau\left(\vartheta\right)=\int_0^\tau f\left(t\right)^2\left[h\left(t\right)+\vartheta g\left(t\right)\right]\mathrm{d}t$$

and

$$\vartheta=\left(\Psi_\tau\left(\vartheta\right)-\int_0^\tau f\left(t\right)^2h\left(t\right)\mathrm{d}t\right)\left(\int_0^\tau f\left(t\right)^2g\left(t\right)\mathrm{d}t\right)^{-1}.$$

Hence the SE is

$$\check\vartheta_{\tau,\varepsilon}=\left(\Psi_{\tau,\varepsilon}-\int_0^\tau f\left(t\right)^2h\left(t\right){\rm d}t\right)\left(\int_0^\tau f\left(t\right)^2g\left(t\right){\rm d}t\right)^{-1}.$$

This estimator is consistent and has the rate of convergence $\varepsilon^{1/3}$.

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

Suppose that we have slightly different partially observed system

$$\begin{aligned}{\rm d}Y_t&=a\left(\vartheta,t\right)Y_t{\rm d}t+b\left(\vartheta,t\right){\rm d}V_t, \qquad Y_0=0,\\ {\rm d}X_t&=f\left(\vartheta,t\right)Y_t{\rm d}t+\varepsilon\sigma\left(t\right){\rm d}W_t, \qquad X_0=0.\end{aligned}$$

As before the process $X^T=(X_t,0\leq t\leq T)$ is observable and Y^T is hidden. All functions $a\left(\cdot\right),b\left(\cdot\right),f\left(\cdot\right)$ and $\sigma\left(\cdot\right)$ are supposed to be known. The parameter $\vartheta\in\Theta=(\alpha,\beta)$ is unknown and has to be estimated by observations X^T.

One way is to use the MLE $\hat\vartheta_{\tau,\varepsilon}$. Remind that the MLE is defined by the equation

$$L(\hat\vartheta_{\tau,\varepsilon},X^\tau)=\sup_{\vartheta\in\Theta}L(\vartheta,X^\tau),$$

where the likelihood ratio function is

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

Here $M\left(\vartheta,t\right)=f\left(\vartheta,t\right)m\left(\vartheta,t\right)$, the function $\sigma\left(\cdot\right)$ is supposed to be positive.

The random process (conditional expectation) $m_t=m\left(\vartheta,t\right)=\mathbf{E}_\vartheta\left(Y_t|\,X_s,0\leq s\leq t\right)$ satisfies the equation of the Kalman-Bucy filtration: $m\left(\vartheta,0\right)=y_0$,

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

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

with initial value $\gamma\left(\vartheta,0\right)=0$. It was shown that this estimator is consistent, asymptotically normal

$$\varepsilon^{-1/2}\left(\hat\vartheta_{\tau,\varepsilon}-\vartheta_0\right)\Longrightarrow\mathcal{N}\left(0,{\rm I}^\tau\left(\vartheta_0\right)^{-1}\right),\ {\rm I}^\tau\left(\vartheta\right)=\int_0^\tau\frac{\dot S\left(\vartheta,t\right)^2}{2S\left(\vartheta,t\right)\sigma\left(t\right)}{\rm d}t$$

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

Introduce notation: $\tau<t\leq T$

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

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

$$\xi_t=\frac{1}{{\rm I}_\tau^t(\vartheta_0)}\int_\tau^t\frac{\dot S\left(\vartheta_0,s\right)}{\sqrt{2S\left(\vartheta_0,s\right)\sigma\left(s\right)}}{\rm d}w\left(s\right).$$

We have to define the random processes

$$M(\check{\vartheta}_{\tau,\varepsilon},s)=f\left(\check{\vartheta}_{\tau,\varepsilon},s\right)m\left(\check{\vartheta}_{\tau,\varepsilon},s\right)$$

and

$$\dot{M}(\check{\vartheta}_{\tau,\varepsilon},s)=\dot{f}\left(\check{\vartheta}_{\tau,\varepsilon},s\right)m\left(\check{\vartheta}_{\tau,\varepsilon},s\right)+f\left(\check{\vartheta}_{\tau,\varepsilon},s\right)\dot{m}\left(\check{\vartheta}_{\tau,\varepsilon},s\right)$$

.

Conditions $\mathcal{C}$.

$\mathcal{C}_1$. *For any $t_0\in(\tau,T]$*

$$\inf_{\vartheta\in\Theta}\mathrm{I}_\tau^{t_0}\left(\vartheta\right)>0.$$

$\mathcal{C}_2$. *The functions $f\left(\cdot\right),\sigma\left(\cdot\right)$ are separated from zero and the functions $f\left(\cdot\right),b\left(\cdot\right)$ have two continuous derivatives w.r.t. ϑ.*

Proposition 1. *Let the conditions $\mathcal{A}$,$\mathcal{B}$,$\mathcal{C}$ be fulfilled then the One-step MLE -process $\vartheta_{t,\varepsilon}^\star,\tau<t\leq T$ is consistent: for any $\nu>0$*

$$\mathbf{P}_{\vartheta_0}\left(\sup_{t_0\leq t\leq T}\left|\vartheta_{t,\varepsilon}^\star-\vartheta_0\right|\geq\nu\right)\longrightarrow0,$$

the random process $\xi_{t,\varepsilon},t_0\leq t\leq T$ converges in distribution in the measurable space $\left(\mathcal{C}\left[t_0,T\right],\mathscr{B}\right])$ to the Gaussian process

$$\frac{\vartheta_{t,\varepsilon}^\star-\vartheta_0}{\sqrt{\varepsilon}}\Longrightarrow\xi_t,\qquad\qquad \xi_t\sim\mathcal{N}\left(0,\mathrm{I}_\tau^t\left(\vartheta_0\right)^{-1}\right)$$

Adaptive filtration.

Note that

$$\frac{m_t-Y_t}{\sqrt{\varepsilon}}=\sqrt{\frac{\sigma\left(t\right)b\left(\vartheta,t\right)}{f\left(\vartheta,t\right)}}\,\zeta_{t,\varepsilon}+o\left(1\right)\Longrightarrow\sqrt{\frac{\sigma\left(t\right)b\left(\vartheta,t\right)}{2f\left(\vartheta,t\right)}}\,\zeta_t,$$

where $\zeta_t,t\in(0,T]$ are independent Gaussian random variables, $\zeta_t\sim\mathcal{N}\left(0,1\right)$.

We have for any, say, continuous function $h\left(t\right)$

$$\varepsilon^{-1/2}\int_0^T(m_t-Y_t)h\left(t\right)\mathrm{d}t\longrightarrow0$$

and

$$\varepsilon^{-1}\int_0^T(m_t-Y_t)^2h\left(t\right)\mathrm{d}t\longrightarrow\int_0^T\frac{\sigma\left(t\right)b\left(\vartheta,t\right)}{2f\left(\vartheta,t\right)h\left(t\right)}\mathrm{d}t.$$

Introduce the equations of adaptive filtration

$$\mathrm{d}\hat{m}\left(t\right)=-q_\varepsilon\left(\vartheta_{t,\varepsilon}^\star,t\right)\hat{m}\left(t\right)\mathrm{d}t+\frac{\gamma\left(\vartheta_{t,\varepsilon}^\star,t\right)f\left(\vartheta_{t,\varepsilon}^\star,t\right)}{\varepsilon^2\sigma\left(t\right)^2}\mathrm{d}X_t,\ \hat{m}\left(0\right)=y_0,$$

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

Here

$$q_\varepsilon\left(\vartheta,t\right)=\frac{\gamma\left(\vartheta,t\right)f\left(\vartheta,t\right)^2}{\varepsilon^2\sigma\left(t\right)^2}-a\left(\vartheta,t\right)$$

and we suppose that Riccati equation can be solved before the experience.

The error of approximation of $m_t = m(\vartheta_0, t)$ is
$$\frac{m_t - \hat m(t)}{\sqrt{\varepsilon}} \Longrightarrow \frac{b(\vartheta_0,t)\,\dot f(\vartheta_0,t)}{\sigma(t)} Y_t - \frac{\dot b(\vartheta_0,t)\sqrt{\sigma(t)}}{\sqrt{f(\vartheta_0,t)\,b(\vartheta_0,t)}}\,\zeta_t\,\xi_t,$$
where $\zeta_t, t\in(0,T]$ are independent standard Gaussian r.v.'s.

This limit together with
$$\frac{m_t - Y_t}{\sqrt{\varepsilon}} \Longrightarrow \sqrt{\frac{\sigma(t)\,b(\vartheta,t)}{2f(\vartheta,t)}}\,\zeta_t,$$
allows us to describe the error of approximation of Y_t by $\hat m(t)$:
$$\frac{\hat m(t) - Y_t}{\sqrt{\varepsilon}}.$$

It is possible to give this couple of equations in recurrent form too
$$\mathrm{d}\tilde m(t) = -q_\varepsilon\left(\vartheta^\star_{t,\varepsilon}, t\right)\tilde m(t)\,\mathrm{d}t + \frac{\hat\gamma(t)\,f\left(\vartheta^\star_{t,\varepsilon},t\right)}{\varepsilon^2\sigma(t)^2}\mathrm{d}X_t, \quad \tilde m(0) = y_0,$$
$$\frac{\partial\hat\gamma(t)}{\partial t} = 2a\left(\vartheta^\star_{t,\varepsilon},t\right)\hat\gamma(t) - \frac{\hat\gamma(t)^2 f\left(\vartheta^\star_{t,\varepsilon},t\right)^2}{\varepsilon^2\sigma(t)^2} + b\left(\vartheta^\star_{t,\varepsilon},t\right)^2,\ \hat\gamma(0)=0.$$
It was shown that $\hat m(t) - m(\vartheta_0,t) = \sqrt{\varepsilon}O_p(1)$.

Remarks.

1. The rate of convergence $\phi_\varepsilon = \varepsilon^m$ of the preliminary estimator for the construction of a consistent One-step MLE-process has to be such that $m\in(\frac14,\frac12]$. Recall that the proposed SE $\check\vartheta_\varepsilon$ has $\phi_\varepsilon = \varepsilon^{1/3}$.

2. It is possible to study this construction supposing that $\tau = \tau_\varepsilon \to 0$ sufficiently slow. We need just to provide for the preliminary estimator $\bar\vartheta_{\tau_\varepsilon,\varepsilon}$ the convergence
$$\mathbf{E}_\vartheta\left|\bar\vartheta_{\tau_\varepsilon,\varepsilon} - \vartheta\right|^2 \le C\varepsilon^{2m}$$
where $m\in(\frac14,\frac12]$. This will allow us to prove the asymptotic efficiency of the One-step MLE-process for all $t\in(0,T]$.

References

[1] Kutoyants Yu.A. On parameter estimation of hidden Ornstein - Uhlenbeck process, // J. Multivariate Analysis. 2019. V. 169. No. 1. P. 248–263.

[2] Kutoyants Yu.A. Quadratic variation estimation of hidden Markov process and related problems, 2022, submitted.

Кутоянц Ю.А. (Университет Ле-Мана, Ле-Ман, Франция, Томский государственный университет, Томск, Россия, 2022) **Статистика скрытых марковских процессов (непрерывное время)**

Аннотация. Представляем обзор нескольких недавних результатов по оцениванию параметров частично наблюдаемых линейных систем и построению адаптивных уравнений фильтрации Калмана-Бьюси. Предполагается, что уравнение наблюдения содержит малый шум уровня ε, а свойства оценок описываются в асимптотике $\varepsilon \to 0$. Адаптивный фильтр строится в несколько этапов. Сначала предлагаем непараметрическую оценку квадратичной вариации производной наблюдений. Затем используем эту оценку для построения одношагового MLE-процесса. Наконец, этот оценочный процесс подставляется в уравнения фильтрации.

Ключевые слова: фильтр Калмана-Бьюси, оценка волатильности, адаптивная фильтрация, одношаговый MLE-процесс.