Estimation and filtering of smooth signals.

1. Kernel type of estimators for the model signal plus WGN.

2. Kalman-type filter approach for this model.

3. On-line estimation for the diffusion observed process.

4. On-line estimation for partially observed signal.

The estimation problem of a function $S(t)$, $t \in [0, T]$ and its derivatives on the base of the observation process $X(t)$ with

$$dX(t) = S(t)\,dt + \varepsilon\,dw(t), \quad X(0) = 0, \tag{1.1}$$

where $w(t)$ is a standard Wiener process, ε is a small parameter, is very popular in the statistical literature. It is enough to mention Wiener [8], who considered this problem assuming that $S(t)$ is a stationary Gaussian process with known covariance function, and Kalman [4], who assumed that $S(t)$ is a solution of a linear stochastic differential equation (SDE), and created Kalman's filter for $S(t)$. In [3] this problem was considered from the non-parametric estimation point of view. It was shown there that the rate of convergence of estimation risks to 0, as $\varepsilon \to 0$, depends on the a priori information concerning smoothness of S.

Denote by $\Sigma(\beta, L)$, $\beta = k + \alpha$, the class of functions S, having k derivatives on $(0, T)$, and, moreover, $S^{(k)}(t)$ is a Hölder continuous with the exponent α function, so that

$$|S^{(k)}(t') - S^{(k)}(t)| \le L|t' - t|^\alpha, \quad 0 < \alpha \le 1.$$

Theorem 1 (IK 80)

Let $g(t)$ be a sufficiently smooth function on R^1 with compact support and

$$\int_{R^1} g(t)\,dt = 1, \quad \int_{R^1} t^i g(t)\,dt = 0, \quad i = 1, 2, \ldots, k.$$

Then the kernel estimators

$$\hat{S}_\varepsilon(t) = \frac{1}{\delta_\varepsilon} \int_{R^1} g\left(\frac{t-u}{\delta_\varepsilon}\right) dX_\varepsilon(u); \quad \hat{S}_\varepsilon^{(j)}(t) = \frac{1}{\delta_\varepsilon^{j+1}} \int_{R^1} g^{(j)}\left(\frac{t-u}{\delta_\varepsilon}\right) dX_\varepsilon(u); \quad j = 1, 2, \ldots, k$$

with $\delta_\varepsilon = \varepsilon^{\frac{2}{2\beta+1}}$ have property

$$\sup_{S \in \Sigma(\beta, L)} E\left\{ \left[\frac{\hat{S}_\varepsilon(t) - S(t)}{\varepsilon^{2\beta/(2\beta+1)}}\right]^2 + \sum_{j=1}^k \left[\frac{\hat{S}_\varepsilon^{(j)}(t) - S^{(j)}(t)}{\varepsilon^{2(\beta-j)/(2\beta+1)}}\right]^2 \right\} \le C, \tag{1.2}$$

and there are no estimators with uniformly in $\Sigma(\beta, L)$ better rate of convergence to 0 risks, as $\varepsilon \to 0$.

Deficiencies of the kernel estimator:

1. $S_\varepsilon(t_0)$ does not help in the estimation of $S(t_0+h)$.
2. Kernels $g(t)$ become more and more cumberse with growth β.

CKL (1997):

On-line estimator was proposed.

$$d\hat{S}_\varepsilon^{(j)}(t) = \hat{S}_\varepsilon^{(j+1)}(t)\,dt$$
$$+ \frac{q_{j+1}}{\varepsilon^{2(j+1)/(2\beta+1)}}(dX^\varepsilon(t) - \hat{S}_\varepsilon^{(0)}(t)\,dt), \quad j=0,1,\ldots,k-1 \quad (1.3)$$
$$d\hat{S}_\varepsilon^{(k)}(t) = \frac{q_{k+1}}{\varepsilon^{2(k+1)/(2\beta+1)}}(dX^\varepsilon(t) - \hat{S}_\varepsilon^{(0)}(t)\,dt), \quad\quad (1.4)$$

subject to the initial conditions $\hat{S}(0) = S_0$, $\hat{S}^{(j)}(0) = S_0^j$, $j=1,\ldots,k$, which reflect *a priori* information on $S(0)$, $S^{(j)}(0)$, $j=1,\ldots,k$.

Denote by $p_k(\lambda)$ the polynomial

$$p_k(\lambda) = \lambda^{k+1} + q_1\lambda^k + \cdots + q_k\lambda + q_{k+1}. \quad\quad (1.5)$$

The following result was proven in ~~[1]~~ **CKL(1997)**

Theorem 2.1. *For any choice $q_1,\ldots,q_{k+1}$ such that all roots of the polynomial $p_k(\lambda)$ have negative real parts, and for arbitrary bounded initial conditions $S_0,\ldots,S_0^k$ the tracking filter (1.4) has the property: there exists an initial boundary layer $\Delta_\varepsilon = C_1\varepsilon^{2/(2\beta+1)}\log(1/\varepsilon)$ such that for $t \geq \Delta_\varepsilon$ the inequality (1.2) is valid.*

The goal of this paper is to widen the result of [1] to the case of diffusion observation process $X_\varepsilon(t)$ with a diffusion coefficient $\varepsilon^2\sigma^2(t,x)$ and a drift coefficient $F(t,x,S)$ depending on unknown but sufficiently smooth function $S(t)$ (for the case $\sigma \equiv 1, F(t,x,S) = S$ we have again the observation process (1.1)). We propose a nonlinear analogy of the filter (1.4) and study the properties of it. Note that the linearity of the filter (1.4) is very essential for its analysis in [1]: the approach in [1] uses explicit solution (in the matrix form) of the system of equations (1.4). Therefore we use here another approach, based on the Lyapunov functions. In order to clarify main idea, this approach is applied initially in Sec. 2 to an alternative proof of Theorem 1.1. Then, in Sec. 3, nonlinear Kalman-type filter is proposed for the nonlinear diffusion observation process, and the optimal rate of convergence of risks to 0 is proven under some assumptions concerning coefficients F and σ, for

Why filter $(1.3),(1.4)$?

$S \in \Sigma(\beta,L)$: auxiliary filtering model.

$$\begin{cases} dS(t) = S^{(1)}(t)\,dt; \\ \quad \cdots \cdots \\ dS^{(k-1)}(t) = S^{(k)}(t)\,dt, \\ dS^{(k)}(t) = \sigma_\varepsilon\, dW_1(t); \\ dX_\varepsilon(t) = S(t)\,dt + \varepsilon\, dW(t) \end{cases}$$

It is possible to check that for $\sigma_\varepsilon = \varepsilon^{\lambda(\beta-k)-1}$ Kalman filt asymptotically, for $\varepsilon \to 0$, converges to the filter $(1.3),(1.4)$ with some $q_1, \ldots, q_{k+1}$.

3

On-line estimation for the
diffusion observed process.

(K.20

$$d X_\varepsilon(t) = F(X_\varepsilon(t), S(t)) dt + \varepsilon \delta (t) dw(t)$$

$$0 \le t \le T.$$

(A1) F, δ is are Lipschitz

continuous w.r.t. x, S, and

δ is bounded.

(A2) For some constants $0 < \alpha_1 < \alpha_2 <$,
the inequalities

$$\alpha_1 \le |F_S'(x, S)| \le \alpha_2$$
hold.

Denote

$$-_\psi V(y) = \sum_{i=0}^{K-1} \frac{\partial V}{\partial y_i} \left(y_{i+1} - \psi \, q_{i+1} y_0 \right) + \frac{\partial V}{\partial y_K} \left(-\psi \, q_{K+1} y_0 \right)$$

(A3) Constants $q_1, \ldots, q_{K+1}$ can be chosen so that $\exists$ quadratic form

$$V(y) = \sum_{i,j=0}^{K} a_{ij} y_i y_j$$

that the inequalities

$$V(y) \geq |y|^2; \quad L_{\frac{\alpha_1}{\alpha_2}} V(y) \leq -\beta |y|^2; \quad L_{\frac{\alpha_2}{\alpha_1}} V(y) \leq -\beta |y|^2$$

hold for some constant $\beta > 0$.

Remark. It is easy to check that (A3) always holds for any $\alpha_1 > 0, \alpha_2 > 0$, if

$$\beta \leq 2 \quad (K = 1 \text{ or } K = 2)$$

Following CKL (1997) consider the on-line estimator for $\left(S^{(0)}(t), \ldots, S^{(K)}(t)\right)$

$$(*) \begin{cases} d\hat{S}_\varepsilon^{(j)}(t) = \hat{S}_\varepsilon^{(j+1)}(t)\, dt + \\ \\ + \dfrac{q_{j+1}}{\gamma^{j+1}} \dfrac{dX_\varepsilon(t) - F(X_\varepsilon(t), \hat{S}_\varepsilon(t))\, dt}{F'_S(X_\varepsilon(t), \hat{S}_\varepsilon(t))}; \\ \\ \hspace{4cm} j = 0, 1, \ldots, K-1, \\ \\ d\hat{S}_\varepsilon^{(K)}(t) = \dfrac{q_{K+1}}{\gamma^{K+1}} \dfrac{dX_\varepsilon(t) - F(X_\varepsilon(t), \hat{S}_\varepsilon(t))\, dt}{F'_S(X_\varepsilon(t), \hat{S}_\varepsilon(t))} \end{cases}$$

subject to the initial conditions

$$\hat{S}_\varepsilon^{(j)}(0) = S_0^j, \quad j = 0, 1, \ldots, K.$$

Here $\gamma = \gamma_\varepsilon = \varepsilon^{\frac{2}{2\beta+1}}$

Theorem 3. Let the conditions $S \in \Sigma(\beta, L)$ and (A1–A3) hold. Then the upper bound

$$\sup_{S \in \Sigma(\beta, L)} E\left\{ \left[\frac{\hat{S}_\varepsilon(t) - S(t)}{\varepsilon^{2\beta/(2\beta+1)}} \right]^2 + \sum_{j=1}^K \left[\frac{\hat{S}_\varepsilon^{(j)}(t) - S^{(j)}(t)}{\varepsilon^{2(\beta-j)/(2\beta+1)}} \right]^2 \right\} \le C$$

is valid for the estimator (8.2)

and for $t > C \, \varepsilon^{\frac{2}{2\beta+1}} \log \frac{1}{\varepsilon}$,

if the constants $q_1, \ldots, q_{K+1}$ are chosen in accordance with (A3),

and $\gamma = \varepsilon^{\frac{2}{2\beta+1}}$.

This rate of convergence risks to 0 cannot be exceeded uniformly in $\Sigma(\beta, L)$ by any other estimator.

Example.

$$dX_\varepsilon(t) = S(t)X_\varepsilon(t)\,dt + \varepsilon\,dw(t)$$

$$X_\varepsilon(0) = 0.$$

$$X_\varepsilon = \varepsilon\,Y_\varepsilon(t)$$

$$dY_\varepsilon(t) = S(t)\,dt + dw(t)$$

Consistent estimator for $S(t)$ does not exist.

On-line estimation of
smooth signal with partial
observations. (C.K.20

$$(.1)\begin{cases} dX_\varepsilon(t) = S(t)\,dt + d\xi_\varepsilon(t,\omega)/\varepsilon \\[4pt] X_\varepsilon(0) = 0 \; ; \; 0 \le t \le T. \\[4pt] \Big(dY_\varepsilon(t) = F(X_\varepsilon(t))\,dt + \varepsilon\sigma\Big(\tfrac{t}{\varepsilon}\Big) \Big)\,dW_{(t)} \\[4pt] Y_\varepsilon(0) = 0. \end{cases}$$

Conditions

B1] F is continuously differ.

and $0 < \alpha_1 < |F'(a)| < \alpha_2$.

B2] $\sigma(t)$ is ~~Lipschitz continu~~

~~and~~ bounded.

(B3) The process $\xi_\varepsilon(t,\omega)$ is independent of w_t^H

and

$$E\,\xi_\varepsilon(t,\omega) = 0$$

$$\sup_{0 \leq t \leq T} E\,|\,\xi_\varepsilon(t,\omega)\,|^2 \leq \alpha_3\,\varepsilon^{2\eta}$$

with

$$\eta \geq \frac{2\beta+2}{2\beta+3}$$

Remark. The condition (B3) holds if

$$\frac{2\beta+2}{2\beta+3} \leq \eta < 1,$$

so the level of noise for unobserved component can be higher under this condition, than the level of noise for the observable component.

Introduce
$$\gamma = \gamma_\varepsilon = \varepsilon^{\frac{2}{2\rho+3}}$$

and consider nonlinear filter
for $\Phi(t) = \int_0^t S(s)\,ds$

$$\begin{cases} d\hat{\Phi}_\varepsilon^{(j)}(t) = \hat{\Phi}_\varepsilon^{(j+1)}(t)\,dt + \\[2mm] \quad + \dfrac{q_{j+1}}{\gamma^{j+1}} \dfrac{d Y_\varepsilon(t) - \hat{\Phi}_\varepsilon^{(0)}(t)\,dt}{F'\left(\hat{\Phi}_\varepsilon^{(0)}(t)\right)} \\[3mm] \qquad\qquad j = 0, \dots, \kappa. \\[4mm] d\hat{\Phi}_\varepsilon^{(\kappa+1)}(t) = \dfrac{q_{\kappa+2}}{\gamma^{\kappa+2}} \dfrac{dY_\varepsilon(t) - \hat{\Phi}_\varepsilon^{(0)}\,dt}{F'(\cdot)} \end{cases}$$

1

Introduce

$$\tilde{L}_r V = \sum_{j=1}^{K} \frac{\partial V}{\partial y_j}(y_{j+1} - q_{j+1} y_0) - \frac{\partial V}{\partial y_{K+1}} \psi q_{K+2} y_0$$

(B4)

The constants $q_1, ..., q_{K+2}$ can be chosen so that

$\exists$ a quadratic form

$V(y_0, y_1, ..., y_{K+1})$ s.t.

$$V(y) \geq |y|^2, \text{ and}$$

$$\tilde{L}_{r_1} V(y) \leq -\beta |y|^2$$

$$\tilde{L}_{r_2} V(y) \leq -\beta |y|^2.$$

Here $\beta > 0$,

$$r_1 = \frac{\alpha_1}{\alpha_2}; \quad r_2 = \frac{\alpha_2}{\alpha_1}.$$

Theorem 4. Assume that the conditions (B1)-(B4) hold for the partial observation model (·), $S \in \Sigma(\beta, L)$ and (

$q_1, \ldots, q_{k+2}$ are chosen in accordance with (B4). Then the nonlinear filter (!) with arbitrary bounded initial conditions and

$$\hat{S}_{\varepsilon}^{(j)}(t) = \hat{\Phi}_{\varepsilon}^{(j+1)}(t) \quad \text{has}$$

properties: for $j = 0, \ldots, k$

and $t > C_0 \, \varepsilon^{\frac{2}{2\beta+3}} \log \frac{1}{\varepsilon}$

$$E \left| \frac{\hat{S}_{\varepsilon}^{(j)}(t) - S^{(j)}(t)}{\varepsilon^{\frac{2(\beta-j)}{2\beta+3}}} \right|^2 < C.$$

Corollary.

Let $S \in \Sigma(\beta, L)$ with $\beta \le 1$, the conditions (B1-B3) hold. Consider partial observations model $(\cdot)$. Let $\hat{\Phi}_\varepsilon^{(0)}(t)$, $\hat{\Phi}_\varepsilon^{(1)}(t)$ solve the filtering equations

$$\begin{cases} d\hat{\Phi}_\varepsilon^{(0)}(t) = \hat{\Phi}_\varepsilon^{(1)}(t)\,dt + \dfrac{q}{\varepsilon^{2/(2\beta+3)}} \dfrac{dY_\varepsilon(t) - F\left(\hat{\Phi}_\varepsilon^{(0)}(t)\right)dt}{F'\left(\hat{\Phi}_\varepsilon^{(0)}(t)\right)} \\[4mm] d\hat{\Phi}_\varepsilon^{(1)}(t) = \dfrac{q}{\varepsilon^{4/(2\beta+3)}} \dfrac{dY_\varepsilon(t) - F\left(\hat{\Phi}_\varepsilon^{(0)}(t)\right)dt}{F'\left(\hat{\Phi}_\varepsilon^{(0)}(t)\right)} \end{cases}$$

subject to the initial conditions

$\hat{\Phi}_\varepsilon^{(0)}(0) = 0$; $\hat{\Phi}_\varepsilon^{(1)}(0) = \Phi^{(1)}(0)$, here $\Phi^{(1)}(0)$ is an arbitrary constant do not depending on ε.

Then for any $q > \dfrac{\alpha_2}{\alpha_1}$, $\exists\, C_0 > 0$ such that for any $t > C_0 \varepsilon^{\frac{2}{2\beta+3}} \log \frac{1}{\varepsilon}$ the upper bound

$$E\left| \hat{\Phi}_\varepsilon^{(1)}(t) - S(t) \right|^2 \le C\, \varepsilon^{\frac{4\beta}{2\beta+3}}$$

holds.

This upper bound is unimprovable in the minimax sense.