Estimation of function from noisy data

V. Solev*

SAPS VII -
March 2009 - Le Mans, France

*Steklov Institute, St Petersbourg, RUSSIA. e-mail: solev@pdmi.ras.ru.

1. Statistical problem

Generally the problem looks as following. On a large time interval $[-T, T]$ we observe a process $Y(t)$,

$$dY(t) = s(t)dt + dX(t), \tag{1}$$

where an unknown function s belongs to a given set $\mathscr{L}_*$,

$$\mathscr{L}_* \subset L^2_{loc},$$

$X(t)$ is a zero-mean gaussian process with stationary increments and the spectral density f. The spectral density f of the noise process is unknown and belongs to a given class of nonnegative functions $\mathscr{K}$.

To estimate an unknown function s one makes observations (for some collection $\mathscr{D}_T$ of smooth functions φ supported on interval $[-T, T]$)

$$\mathbf{y}[\varphi] = \mathbf{s}[\varphi] + \mathbf{x}[\varphi],$$

and constructs an estimator $\hat{s}_T$, based on these observations. Here we set

$$\mathbf{y}[\varphi] = \int\limits_{\mathbb{R}^1} \varphi(t)\, dY(t),$$

and define by the same way variables $\mathbf{s}[\varphi]$ and $\mathbf{x}[\varphi]$ for $\varphi \in \mathscr{D}_T$.

We denote by $\mathscr{L}$ the Banach space of locally square integrable functions with the norm $\|s\|^2_{\mathscr{L}}$:

$$\|s\|^2_{\mathscr{L}} = \sup_x \int\limits_x^{x+1} |s(t)|^2\, dt,$$

and assume that $\mathscr{L}_* \subset \mathscr{L}$.

An estimator $\hat{s}_T(\cdot)$ is $\mathfrak{F}_T-$measurable random element of the space $\mathscr{L}$ such that $\hat{s}_T \in \mathscr{L}_*$. Here the $\sigma-$algebra $\mathfrak{F}_T$ is defined by

$$\mathfrak{F}_T = \sigma\left\{\mathbf{y}[\varphi], \varphi \in \mathscr{D}_T\right\}.$$

The set of all such estimators we denote by $\mathscr{S}(T)$.

Consider as the risk function of an estimator $\widehat{s}_T$ for s

$$R\left(\widehat{s}_T, \mathscr{L}_*\right) = \sup_{s \in \mathscr{L}_*} \mathbf{E}_f \left\| \widehat{s}_T - s \right\|^2_{\mathscr{L}},$$

and minimax risk

$$R\left(T, \mathscr{L}_*\right) = \inf_{\widehat{s}_T \in \mathscr{S}(T)} R\left(\widehat{s}_T, \mathscr{L}_*\right).$$

For a given estimator $\widehat{s} \in \mathscr{S}_T$, we can take the ratio

$$\rho(\widehat{s}_T, T, \mathscr{L}_*, f) = \frac{R\left(\widehat{s}_T, \mathscr{L}_*\right)}{R\left(T, \mathscr{L}_*\right)},$$

in order to compare an estimator $\widehat{s}_T$ and the minimax estimator. Our goal is to construct an suboptimal estimator $\widehat{s}_T$ such that, for sufficiently large T

$$\rho(\widehat{s}_T, T, \mathscr{L}_*, f) \leq C(\mathscr{K}, \mathscr{L}_*).$$

2. Parametric set $\mathscr{L}_*$

For a countable set $\Lambda \subset \mathbb{R}^1$ (which will be called spectral set) denote

$$\kappa = \kappa\,(\Lambda) = \inf_{u,v\in\Lambda, u\neq v} |u - v| > 0.$$

Consider Stepanov class $\mathscr{L}(\Lambda)$ locally square integrable functions

$$s(t) = \sum_{u\in\Lambda} a(u)e^{itu}, \; \sum_{u\in\Lambda} |a(u)|^2 < \infty.$$

Paley R. and Wiener N. proved, that under the condition $\kappa\,(\Lambda) > 0$

$$C_1 \sum_{u\in\Lambda} |a(u)|^2 \leq \|s\|^2_{\mathscr{L}} = \sup_x \int\limits_x^{x+1} |s(t)|^2\, dt \leq C_2 \sum_{u\in\Lambda} |a(u)|^2$$

where C_1, C_2 depend only on κ.

It may be proved that the Banach norm $||\cdot||_{\mathscr{L}}$ is topologically equivalent on $\mathscr{L}(\Lambda)$ to the Hilbert norm $||\cdot||_{L^2(-T,T)}$,

$$||s||^2_{L^2(-T,T)} = \frac{1}{2T}\int_{-T}^{T} |s(t)|^2\,dt$$

for sufficiently large $T: \; T > T_0(\kappa)$.

As the parametric set for unknown function s we consider the subset $\mathscr{L}_*$ of the space $\mathscr{L}$ defined by

$$s(t) = \sum_{u\in\Lambda} a(u)e^{itu}, \; \sum_{u\in\Lambda} |a(u)|^2(1+|u|)^{2\beta} < C.$$

Let $\beta = r + \alpha$, where $r > 0$ is an integer, and $\alpha \in (0,1)$. For analytical goal it is convenient to use seminorm

$$\|s\|_\beta = \int\limits_{-\infty}^{\infty} \frac{\left\|s^{(r)}(t+y) - s^{(r)}(t)\right\|^2_{L^2_{[-T,\,T]}}}{|y|^{1+2\alpha}}\,dy$$

3. Class $\mathscr{K}$ of spectral densities

We define the class $\mathscr{K} = \mathscr{K}(K)$ of nonnegative functions by

$$\lambda(f) = \sup_I \frac{1}{|I|} \int_I f(u)\, du \times \frac{1}{|I|} \int_I \frac{1}{f(u)}\, du \le K < \infty, \ f \in \mathscr{K}.$$

Here supremum is taken over all intervals I.

We try to explain the choice of the class $\mathscr{K}$. Suppose we observe on interval $t \in [-T, T]$ unknown function

$$s(t) = \sum_{u \in \Lambda} a(u) \varphi_u(t),$$

in stationary noise:

$$dY(t) = s(t)dt + dX(t)$$

with spectral density f. In our case

$$\varphi_u(t) = e^{itu}.$$

The reasonable estimator $\widehat{a}_T(u)$ for unknown coefficient $a(u)$ is defined by

$$\widehat{a}_T(u) = \frac{1}{2T}\int\limits_{-T}^{T} \overline{\psi_u^T(t)}\, dY(t) = a(u) + \frac{1}{2T}\int\limits_{-T}^{T} \overline{\psi_u^T(t)}\, dX(t).$$

Here $\{\psi_u^T(t),\ u \in \Lambda\}$ – is the system of conjugate function,

$$\frac{1}{2T}\int\limits_{-T}^{T} \overline{\psi_u^T(t)}\varphi_v(t)\, dt = \delta_{u,v}.$$

So, we have observations

$$\widehat{a}_T(u) = a(u) + X_u,\ u \in \Lambda,$$

where $X_u, u \in \Lambda$ is zero mean gaussian process, and we have to estimate vector $(a(u), u \in \Lambda)$. The key point for our choice of the class $\mathscr{K}$ of spectral densities is the following result.

Proposition 1. Suppose $f \in \mathscr{K}$, $\kappa > 0$, then

$$\mathbf{E}\left(X_u - \sum_{v \neq u} b(v) X_v\right)^2 \geq C(\kappa, K)\mathbf{E}X_u^2.$$

Now we take a function $\varphi(t), \sup\varphi \subset [-T, T]$,

$$\int_{-T}^{T} \overline{\varphi(t)}\varphi_u(t)\, dt = 0,\ u \in \Lambda \tag{$*$}$$

and consider the observation

$$\xi = \int_{-T}^{T} \overline{\psi_u^T(t)}\, dY(t) = \int_{-T}^{T} \overline{\psi_u^T(t)}\, dX(t) = X[\varphi].$$

We add the observation ξ to the system $\{\widehat{a}_T(u),\ u \in \Lambda\}$. In the case as ξ and $\{\widehat{a}_T(u),\ u \in \Lambda\}$ are independent we have not any new information about unknown vector $(a(u), u \in \Lambda)$.

Denote

$$H_\Lambda = \overline{\mathrm{sp}}\left\{X_u,\ u \in \Lambda\right\},\ H_0 = \overline{\mathrm{sp}}\left\{X[\varphi], \mathrm{sup}\varphi \subset [-T, T], \varphi \text{ satisfies } (*)\right\}$$

In order to control minimax risk in general case, when we use for estimating only vector $(\widehat{a}_T(u),\ u \in \Lambda)$, we need to know, that the unit ball of H_Λ separated from the unit ball of H_0.

Proposition 1. Suppose $f \in \mathscr{K}$, $\kappa > 0$, then there exists constant $C(K, \kappa) > 0$ such that, for $X \in H_\Lambda$ and $\xi \in H_0$,

$$\mathbf{E}\,(X - \xi)^2 \geq C(K, \kappa)\mathbf{E}X^2.$$

4. Suboptimal estimator

At the beginning we consider an example as we observe

$$y = \theta + X,\ X \in N(0, \sigma^2),\ |\theta| \le \tau.$$

In this case the risk R_L of linear estimator

$$\widehat{\theta} = y \frac{\tau^2}{\tau^2 + \sigma^2}$$

may be calculated, and

$$R_L = \frac{\tau^2 \sigma^2}{\tau^2 + \sigma^2}.$$

It is well known (Ibragimov, Hasminskii), that there exist absolute constant μ such that, for minimax risk R,

$$R_L \le \mu R.$$

Now we take the estimator

$$\hat{\theta} = \begin{cases} y, & |y| \leq \tau \\ \tau, & \text{else.} \end{cases}$$

Risk $R(\widehat{\theta})$ of this estimator satisfies to

$$R(\widehat{\theta}) \leq 8 \frac{\tau^2 \sigma^2}{\tau^2 + \sigma^2}.$$

Now consider our problem. Suppose we have observation

$$\widehat{a}_T(u) = a(u) + \frac{1}{2T} \int_{-T}^{T} \overline{\psi_u^T(t)}\, dX(t), \; u \in \Lambda,$$

and the parametric set is defined by

$$\sum_{u \in \Lambda} |a(u)|^2 \left(1 + |u|\right)^{2\beta} \leq C.$$

Let R be the smallest positive value such that

$$\sum_{u\in\Lambda,\,|u|\le R} |\widehat{a}_T(u)|^2 \left(1+|u|\right)^{2\beta} \ge C.$$

We take as estimator $\widehat{s}_T$ for s the function

$$\widehat{s}_T = \sum_{u\in\Lambda,\,|u|\le R} \widehat{a}_T(u)e^{iut}. \qquad **$$

Theorem. Suppose that $\kappa > 0$, and

$$\lambda(f) = \sup_I \frac{1}{|I|}\int_I f(u)\,du \times \frac{1}{|I|}\int_I \frac{1}{f(u)}\,du \le K < \infty,$$

then there exists constant $C(\kappa, K)$ such that for sufficiently large T

$$R\left(\widehat{s}_T, \mathscr{L}_*\right) \le C(\kappa, K) R\left(T, \mathscr{L}_*\right).$$