Weierstraß–Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Maximum likelihood estimate for nonparametric signal in white noise by optimal control

G.N. Milstein[1] M. Nussbaum[2]

submitted: 12th July 2000

[1] Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstr. 39
D-10117 Berlin, Germany
E-Mail: milstein@wias-berlin.de

and

Ural State University
Lenin Street 51
620083 Ekaterinburg
Russia

[2] Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstr. 39
D-10117 Berlin, Germany

and

Department of Mathematics
Cornell University
Malott Hall
Ithaca NY 14853-4201, USA
E-Mail: mn66@cornell.edu

Preprint No. 596
Berlin 2000

2000 *Mathematics Subject Classification.* 62G05, 60H10, 49K15.

Key words and phrases. Nonparametric diffusion model, maximum likelihood method, optimal estimation.

Edited by
Weierstraß–Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
D — 10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail (X.400): c=de;a=d400-gw;p=WIAS-BERLIN;s=preprint
E-Mail (Internet): preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract. The paper is devoted to questions of constructing the maximum likelihood estimate for a nonparametric signal in white noise by considering corresponding problems of optimal control. For signals with bounded derivatives, sensitivity theorems are proved. The theorems state a stability of the maximum likelihood estimate with respect to changing output data. They make possible to reduce the original problem to a standard problem of optimal control which is solved by iterative procedure. For signals of Sobolev type the maximum likelihood estimate is obtained to within a parameter which can be found from a transcendental equation.

1. Introduction

Let us consider the model [4]

$$da(t) = x(t)dt + \varepsilon dw(t),\ 0 \le t \le 1, \quad a(0) = 0, \tag{1.1}$$

where $a(t)$ is an observation, $x(t)$ is an unknown signal, $w(t)$ is a standard Wiener process, $\varepsilon > 0$ is a small parameter.

The vast literature is devoted to different aspects connected with this model. We are dealing here with maximum likelihood estimate of the unknown signal. Let ν be a measure in the space $\mathbf{C}[0,1]$ which is generated by the process $w(t)$. Then the likelihood function [4] is equal to

$$\frac{dP(a(\cdot)/\varepsilon)}{d\nu} = \exp\left\{\frac{1}{\varepsilon^2}\int_0^1 x(t)da(t) - \frac{1}{2\varepsilon^2}\int_0^1 x^2(t)dt\right\}.$$

When it is known that $x(\cdot)$ belongs to a class $\mathbf{K}$, then the maximum likelihood method of finding an estimate for $x(\cdot)$ leads to the problem

$$I = \frac{1}{2}\int_0^1 x^2(t)dt - \int_0^1 x(t)da(t) \longrightarrow \min_{x(\cdot)\in\mathbf{K}}. \tag{1.2}$$

A profound theoretical investigation of the problem is done in [7]. Necessary and sufficient conditions for existence, uniqueness and consistency of the maximum likelihood estimator are given there. The conditions are formulated in terms of some characteristics of the class $\mathbf{K}$. In [7] a number of properties of the maximum likelihood estimator $\hat{x}(a(\cdot))$ are considered as well. For example, a measure of that $a(\cdot)$ for which $\|\hat{x}(a(\cdot)) - x\| \ge r$ is studied. At the same time methods of constructing $\hat{x}(a(\cdot))$ are not considered in full measure up to now. Apparently, to this aim one can apply, for instance, the approach of [5], [6] after a suitable discretization of model (1.1). However we prefer to give a direct solution of problem (1.2) for some important classes $\mathbf{K}$.

Suppose it is known that each function $x(\cdot)$ of a class $\mathbf{K}$ has a derivative $x'(\cdot)$ which is in $\mathbf{L}_2[0,1]$. In this case the functional (1.2) transforms

$$I = \frac{1}{2}\int_0^1 x^2(t)dt + \int_0^1 a(t)x'(t)dt - a(1)x(1), \tag{1.3}$$

and problem (1.2) amounts to a problem of optimal control

$$(1.4) \qquad I = \int_0^1 (\frac{1}{2}x^2 + a(t)u)dt - a(1)x(1) \longrightarrow \min,$$

$$(1.5) \qquad x' = u,$$

in which restrictions to control u and phase variable x are connected with the class $\mathbf{K}$.

Here we consider such problems for two types of a prior information concerning the unknown signal $x(\cdot)$. If the signal $x(\cdot)$ is of Sobolev's type, we treat the class $\mathbf{K}$ of the form

$$(1.6) \qquad \mathbf{K} = \left\{x(\cdot) : \exists x'(\cdot) = u(\cdot) \in \mathbf{L}_2[0,1]\ , \right.$$
$$\left. \frac{1}{2}\int_0^1 (\alpha x^2(t) + u^2(t))dt \le M\ , \quad \alpha \ge 0,\ M > 0 \right\}\ ,$$

where α and M are known constants. In this class and in other Sobolev's classes it is possible to obtain the maximum likelihood estimate $\hat{x}(t)$ for signal $x(t)$ to within a parameter which can be found from a transcendental equation (see Section 7).

If the signal $x(\cdot)$ has a bounded derivative of the order n, we treat the class $\mathbf{K}_n$ of the form

$$(1.7) \qquad \mathbf{K}_n = \left\{x(\cdot) : \exists\ x^{(n-1)}(t) \text{ which is an absolutely continuous function,} \right.$$
$$\left. |x^{(n)}(t)| \le M_n,\ M_n > 0 \right\},$$

where M_n is a known constant. For class (1.7), we consider the more convenient problem than (1.4)-(1.5). To this end we replace the output data $a(t)$ which have bad analytical properties with a little modified data $\bar{a}(t)$ such that there exists a piecewise continuous derivative $\bar{a}'(t)$. In Section 2 we show that the processing with a little change output data gives results closing to optimal ones. The results of Section 2 have not only the subsidiary but also an independent sense. They state the stability of the maximum likelihood method with respect to changing output data. After replacement $a(t)$ by $\bar{a}(t)$ the problem (1.4)-(1.5) can be reduced to the following problem

$$(1.8) \qquad I = \frac{1}{2}\int_0^1 (x(t) - \bar{a}'(t))^2 dt \longrightarrow \min_{x(\cdot)\in\mathbf{K}_n}\ .$$

The problem (1.8) is a fairly known problem and has already been investigated by methods of optimal control in [2] and [3]. We give a detailed presentation for the case $n = 1$ in Sections 3-5 and some generalizations of the discussed problems in Section 6.

2. Sensitivity theorems for signals with bounded derivative

At the beginning let us consider the class of functions (see (1.7))

$$(2.1) \qquad \mathbf{K}_1 = \{x(\cdot) : x(t) \text{ is absolutely continuous, } |x'(t)| \le M,\ M > 0\}$$

and the minimization problem in this class

$$I=\frac{1}{2}\int_0^1 x^2(t)dt+\int_0^1 a(t)x'(t)dt-a(1)x(1)\longrightarrow \min_{x(\cdot)\in\mathbf{K}_1}. \tag{2.2}$$

It is possible to prove that there exists a solution of the problem.

Theorem 2.1. *Let* $\bar{a}(t)$ *be a continuous function such that*

$$\bar{a}(0)=0,\ \bar{a}(1)=a(1), \tag{2.3}$$

$$\int_0^1 |\bar{a}(s)-a(s)|ds\le\delta. \tag{2.4}$$

Let $x_0(\cdot)$ *be a solution of the minimization problem* (2.2) *and* $\bar{x}_0(\cdot)$ *be a solution of the following minimization problem in the same class*

$$\bar{I}=\frac{1}{2}\int_0^1 x^2(t)dt+\int_0^1 \bar{a}(t)x'(t)dt-\bar{a}(1)\cdot x(1)\longrightarrow \min_{x(\cdot)\in\mathbf{K}_1}. \tag{2.5}$$

Then

$$0\le I(\bar{x}_0(\cdot))-I(x_0(\cdot))\le 2\delta M, \tag{2.6}$$

$$\int_0^1 (\bar{x}_0(t)-x_0(t))^2dt\le 4\delta M, \tag{2.7}$$

and if $\delta\le M/3$,

$$\max_{0\le t\le 1}|\bar{x}_0(t)-x_0(t)|\le(24\delta M^2)^{\frac{1}{3}}. \tag{2.8}$$

Proof. Obviously

$$I(\bar{x}_0(\cdot))-I(x_0(\cdot))\ge 0,$$
$$\bar{I}(x_0(\cdot))-\bar{I}(\bar{x}_0(\cdot))\ge 0.$$

Furthermore

$$|I(\bar{x}_0(\cdot))-\bar{I}(\bar{x}_0(\cdot))|=|\int_0^1(a(t)-\bar{a}(t))\bar{x}_0'(t)dt|\le M\delta.$$

Analogously

$$|\bar{I}(x_0(\cdot))-I(x_0(\cdot))|\le M\delta.$$

Therefore

$$\begin{aligned}0\le I(\bar{x}_0(\cdot))-I(x_0(\cdot))&\le I(\bar{x}_0(\cdot))-I(x_0(\cdot))+\bar{I}(x_0(\cdot))-\bar{I}(\bar{x}_0(\cdot))\\&\le|I(\bar{x}_0(\cdot))-\bar{I}(\bar{x}_0(\cdot))|+|\bar{I}(x_0(\cdot))-I(x_0(\cdot))|\le 2\delta M.\end{aligned}$$

Thus the inequality (2.6) is proved.

For derivation of (2.7), let us note that

$$(1-\alpha)x_0(\cdot)+\alpha\bar{x}_0(\cdot)\in\mathbf{K}_1,\ 0\le\alpha\le 1,$$

and introduce the function $f(\alpha)$, $0 \le \alpha \le 1$, (see (1.2))

$$f(\alpha) = I((1-\alpha)x_0(\cdot) + \alpha\bar{x}_0(\cdot))$$

$$= \frac{1}{2}\int_0^1 ((1-\alpha)x_0(s) + \alpha\bar{x}_0(s))^2 ds - \int_0^1 ((1-\alpha)x_0(s) + \alpha\bar{x}_0(s))da(s),$$

which is a quadratic trinomial on α.

Obviously

$$f(0) = I(x_0(\cdot)) \le f(\alpha),$$

and therefore

$$f'(0) \ge 0.$$

We have

$$f'(\alpha) = \int_0^1 ((1-\alpha)x_0(s) + \alpha\bar{x}_0(s))(\bar{x}_0(s) - x_0(s))ds - \int_0^1 (\bar{x}_0(s) - x_0(s))da(s),$$

$$f''(\alpha) = \int_0^1 (\bar{x}_0(s) - x_0(s))^2 ds = const = C > 0.$$

Further,

$$f'(\alpha) = f'(0) + \int_0^\alpha f''(\alpha)d\alpha = f'(0) + C\alpha,$$

$$f(\alpha) = f(0) + \int_0^\alpha (f'(0) + C\alpha)d\alpha = I(x_0(\cdot)) + f'(0)\alpha + C\frac{\alpha^2}{2},$$

and

$$f(1) = I(\bar{x}_0(\cdot)) = I(x_0(\cdot)) + f'(0) + \frac{C}{2}.$$

From here and (2.6)

$$f'(0) + \frac{C}{2} \le 2\delta M$$

and, as $f'(0) \ge 0$, we obtain the inequality (2.7).

Now prove the inequality (2.8). Let

$$m = \max_{0 \le t \le 1} |\bar{x}_0(t) - x_0(t)| = |\bar{x}_0(t^*) - x_0(t^*)|.$$

For certainty we take

$$x_0(t^*) = x_0^* < \bar{x}_0(t^*) = x_0^* + m.$$

Since $|x_0'(t)| \le M$ and $|\bar{x}_0'(t)| \le M$ for $0 \le t \le 1$, it is clear that for $0 \le t \le t^*$

$$x_0(t) \le x_0^* - M(t-t^*),\ \bar{x}_0(t) \ge x_0^* + m + M(t-t^*)$$

and for $t^* \le t \le 1$

$$x_0(t) \le x_0^* + M(t-t^*),\ \bar{x}_0(t) \ge x_0^* + m - M(t-t^*).$$

Hence

$$4\delta M \geq \int_0^1 (\bar{x}_0(s) - x_0(s))^2 ds \tag{2.9}$$

$$\geq \int_{0\vee(t^*-\frac{m}{2M})}^{t^*} (m + 2M(s-t^*))^2 ds + \int_{t^*}^{1\wedge(t^*+\frac{m}{2M})} (m - 2M(s-t^*))^2 ds.$$

We have to find the largest m for which this inequality can take place. Clearly one can seek required m from (2.9) at $t^* = 0$.

We have

$$4\delta M \geq \int_0^{1\wedge\frac{m}{2M}} (m - 2Ms)^2 ds = \begin{cases} m^3/6M, \ m/2M \leq 1, \\ \\ m(m-2M) + 4M^2/3, \ m/2M > 1. \end{cases} \tag{2.10}$$

But for $\delta \leq M/3$ the second case in (2.10) is impossible and therefore

$$m^3 \leq 24\delta M^2.$$

Theorem 2.1 is proved.

Remark 2.1. It is possible to avoid the condition $\bar{a}(1) = a(1)$ in (2.3). To this end let us obtain a prior bound for $|x_0(1)|$.

We have (in (2.11) $x_0(1)$ is denoted by x_0^1 and for definiteness $x_0^1 > 0$)

$$\frac{1}{2}\int_0^1 x_0^2(t)dt \geq \tag{2.11}$$

$$\geq \begin{cases} \frac{1}{2}\int_{(M-x_0^1)/M}^1 (x_0^1 + M(t-1))^2 dt = (x_0^1)^3/6M, \ x_0^1 < M, \\ \\ \frac{1}{2}\int_0^1 (x_0^1 + M(t-1))^2 dt = \frac{1}{2}((x_0^1)^2 - x_0^1 M + M^2/3), \ x_0^1 \geq M. \end{cases}$$

Introduce the function $\varphi(x_0(1))$ (see the right-hand part of (2.11))

$$\varphi(x_0(1)) = \begin{cases} |x_0(1)|^3/6M, \ |x_0(1)| < M, \\ \\ \frac{1}{2}(|(x_0(1)|^2 - |x_0(1)|M + M^2/3), \ |x_0(1)| \geq M, \end{cases}$$

by which we can rewrite the relation (2.11) for all $x_0(1)$ as

$$\frac{1}{2}\int_0^1 x_0^2(t)dt \geq \varphi(x_0(1)).$$

Therefore

$$0 \geq I((x_0(\cdot)) \geq \varphi(x_0(1)) - |a(1)||x_0(1)| - M\int_0^1 |a(t)|dt.$$

Clearly one can take the only positive root X^* of the following equation

$$\varphi(X) - |a(1)|X - M\int_0^1 |a(t)|dt = 0$$

as an upper bound for $|x_0(1)|$. A simple but rough bound for X^* gives the prior bound for $|x_0(1)|$:

$$|x_0(1)| \le X^* \le M + 2|a(1)| + \sqrt{3}\int_0^1 |a(t)|dt. \tag{2.12}$$

Give another derivation of a prior bound for $|x_0(1)|$ which will be useful below. We have

$$\frac{1}{2}\int_0^1 x_0^2(t)dt = \frac{1}{2}\int_0^1 (x_0(1) + \int_1^t x_0'(s)ds)^2 dt$$
$$\ge \frac{1}{2}\int_0^1 (\frac{1}{2}(x_0(1))^2 - (\int_1^t x_0'(s)ds)^2)dt$$
$$\ge \frac{1}{4}|x_0(1)|^2 - \frac{1}{2}\int_0^1 M^2|t-1|^2 dt = \frac{1}{4}|x_0(1)|^2 - \frac{1}{6}M^2.$$

Here we use a simple inequality $\frac{1}{2}(a+b)^2 \le a^2 + b^2$, where

$$a = x_0(1) + \int_1^t x_0'(s)ds,\ b = -\int_1^t x_0'(s)ds.$$

Consequently

$$0 \ge I((x_0(\cdot)) \ge \frac{1}{4}|x_0(1)|^2 - \frac{1}{6}M^2 - |a(1)||x_0(1)| - M\int_0^1 |a(t)|dt.$$

Therefore

$$|x_0(1)| \le 2|a(1)| + 2(|a(1)|^2 + M\int_0^1 |a(t)|dt + \frac{1}{6}M^2)^{1/2} \tag{2.13}$$

and we obtain a new kind of the prior bound for $|x_0(1)|$.

Replace now conditions (2.3) and (2.4) by

$$\bar{a}(0) = 0,\ |\bar{a}(1) - a(1)| \le \delta_1, \tag{2.14}$$

$$\int_0^1 |\bar{a}(s) - a(s)|ds \le \delta_2. \tag{2.15}$$

Similar to (2.12) (or (2.13)) we have for $|\bar{x}_0(1)|$:

$$|\bar{x}_0(1)| \le \bar{X}^*.$$

Clearly $\bar{X}^*$ is close to X^*.

Let δ be such that

$$2M\delta_2 + (X^* + \bar{X}^*)\delta_1 \le 2M\delta.$$

Now we need in small changes of the proof of Theorem 2.1 for affirming (2.6). The proof of (2.7), (2.8) remains without any changes.

Thus all the conclusions of Theorem 2.1 under conditions (2.14)-(2.15) are valid.

Consider the class of functions

$$\mathbf{K}_2 = \{x(\cdot) : x'(t) \text{ is absolutely continuous and } |x''(t)| \le M_2,\ M_2 > 0\}. \tag{2.16}$$

The functional (1.2) in the class $\mathbf{K}_2$ can be rewritten as

$$I(x(\cdot)) = \frac{1}{2}\int_0^1 x^2(t)dt - \int_0^1 (\int_0^t a(s)ds)x''(t)dt + x'(1)\int_0^1 a(s)ds - a(1)x(1). \tag{2.17}$$

It is possible to prove that there exists a solution of the minimization problem for the functional (2.17) in the class $\mathbf{K}_2$.

Theorem 2.2. *Let $\bar{a}(t)$ be a continuous function such that*

$$\bar{a}(0) = 0,\ \bar{a}(1) = a(1), \tag{2.18}$$

$$\int_0^1 \bar{a}(s)ds = \int_0^1 a(s)ds, \tag{2.19}$$

$$\int_0^1 |\int_0^t \bar{a}(s)ds - \int_0^t a(s)ds|dt \le \delta. \tag{2.20}$$

Let $x_0(\cdot)$ be a solution of the minimization problem for the functional (2.17) in the class $\mathbf{K}_2$ and let $\bar{x}_0(\cdot)$ be a solution of the following minimization problem in the same class

$$\bar{I} = \frac{1}{2}\int_0^1 x^2(t)dt - \int_0^1 (\int_0^t \bar{a}(s)ds)x''(t)dt + x'(1)\int_0^1 \bar{a}(s)ds - \bar{a}(1)x(1) \longrightarrow \min_{x(\cdot)\in\mathbf{K}_2}. \tag{2.21}$$

Then

$$0 \le I(\bar{x}_0(\cdot)) - I(x_0(\cdot)) \le 2\delta M_2, \tag{2.22}$$

$$\int_0^1 (\bar{x}_0(t) - x_0(t))^2 dt \le 4\delta M_2, \tag{2.23}$$

$$|x_0(t)| \le M_0,\ |\bar{x}_0(t)| \le M_0,\ |x_0'(t)| \le M_1,\ |\bar{x}_0'(t)| \le M_1,\ 0 \le t \le 1, \tag{2.24}$$

where M_0 and M_1 depend only on $|a(1)|$, $|\int_0^1 a(s)ds|$, $\int_0^1 |\int_0^t a(s)ds|dt$, M_2, and δ.

Then there exists a constant $K > 0$ such that if $\delta \le K$, then

$$\max_{0\le t\le 1} |\bar{x}_0(t) - x_0(t)| \le K_0\delta^{1/3}, \tag{2.25}$$

(2.26) $$\int_0^1 (\bar{x}_0'(t) - x_0'(t))^2 dt \le K_1 \delta^{1/2},$$

(2.27) $$\max_{0\le t\le 1} |\bar{x}_0'(t) - x_0'(t)| \le K_2 \delta^{1/6},$$

where K, K_0, K_1, K_2 depend on M_1 and M_2 only.

Proof. The inequalities (2.22) and (2.23) can be obtained without any essential modifications in compare with the proof of (2.6) and (2.7).

Similar to proving (2.13) we can write

$$\frac{1}{2}\int_0^1 x_0^2(t)dt = \frac{1}{2}\int_0^1 (x_0(1) + x_0'(1)(t-1) + \int_1^t (\int_1^s x_0''(s_1)ds_1)ds)^2 dt$$
$$\ge \frac{1}{4}\int_0^1 (x_0(1) + x_0'(1)(t-1))^2 dt - \frac{1}{2}\int_0^1 (\int_1^t (\int_1^s x_0''(s_1)ds_1)ds)^2 dt$$
$$\ge \frac{1}{4}((x_0(1))^2 - x_0(1)x_0'(1) + \frac{1}{3}(x_0'(1))^2) - \frac{M_2^2}{40}.$$

Further, from representation (2.17) we get

$$0 \ge I(x_0(\cdot)) \ge \frac{1}{4}(|(x_0(1)|^2 - |x_0(1)||x_0'(1)| + \frac{1}{3}|x_0'(1)|^2) - \frac{M_2^2}{40}$$
$$-|a(1)||(x_0(1)| - |\int_0^1 a(s)ds||x_0'(1)| - M_2\int_0^1 |\int_0^t a(s)ds|dt.$$

Since the expression $(|(x_0(1)|^2 - |x_0(1)||x_0'(1)| + \frac{1}{3}|x_0'(1)|^2)$ is a positive definite quadratic form with respect to $|x_0(1)|$ and $|x_0'(1)|$, we obtain from here that $|x_0(1)|$ and $|x_0'(1)|$ are bounded and their bounds depend only on $|a(1)|$, $|\int_0^1 a(s)ds|$, $\int_0^1 |\int_0^t a(s)ds|dt$, and M_2. The same is also true for $|\bar{x}_0(1)|$ and $|\bar{x}_0'(1)|$. Let us note in passing that if $\bar{a}(1)$, and $\int_0^1 \bar{a}(s)ds$ are close respectively to $a(1)$, and $\int_0^1 a(s)ds$ and if δ is small, then the bounds for $|\bar{x}_0(1)|$ and $|\bar{x}_0'(1)|$ are close to ones for $|x_0(1)|$ and $|x_0'(1)|$. Clearly the inequalities (2.24) are a simple consequence of conditions $|x_0''(t)| \le M_2$, $|\bar{x}_0''(t)| \le M_2$.

For (2.25) it is sufficient to mark that instead of (2.10) it can be easily obtained the following inequality

(2.28) $$4\delta M_2 \ge \begin{cases} m^3/6M_1, \ m/2M_1 \le 1, \\ m(m-2M_1) + 4M_1^2/3, \ m/2M_1 \ge 1, \end{cases}$$

and for $\delta \le M_1^2/3M_2$ the second case in (2.28) is impossible (as K_0 in (2.25) one can take $(24M_1M_2)^{1/3}$).

Now let us proceed to the derivation of (2.26) and (2.27). We have

$$(2.29)\quad \int_0^1 (\bar{x}_0'(t)-x_0'(t))^2 dt = (\bar{x}_0'(1)-x_0'(1))(\bar{x}_0(1)-x_0(1))-(\bar{x}_0'(0)-x_0'(0))(\bar{x}_0(0)-x_0(0))$$

$$-\int_0^1 (\bar{x}_0(t)-x_0(t))(\bar{x}_0''(t)-x_0''(t))dt \le 2K_0\delta^{1/3}m_1 + 2M_2(4\delta M_2)^{1/2},$$

where we introduce m_1 as

$$m_1 = \max_{0\le t\le 1} |\bar{x}_0'(t)-x_0'(t)|.$$

In just the same way as we have derived (2.10) we obtain

$$(2.30)\qquad 2K_0\delta^{1/3}m_1 + 4M_2^{3/2}\delta^{1/2} \ge \int_0^1 (\bar{x}_0'(t)-x_0'(t))^2 dt$$

$$\ge \begin{cases} m_1^3/6M_2, & m_1/2M_2 \le 1, \\ m_1(m_1-2M_2)+4M_2^2/3, & m_1/2M_2 \ge 1. \end{cases}$$

By means of decrease of K we can exclude the second case in (2.30) as before and get the inequality

$$2K_0\delta^{1/3}m_1 + 4M_2^{3/2}\delta^{1/2} \ge m_1^3/6M_2.$$

From here and (2.29) it is not difficult to obtain (2.26) and (2.27). Theorem 2.2 is proved.

Remark 2.2. We do not try to obtain any exact bounds. Our principal aim is to show that the processing with a little changed output data gives results closing to optimal ones. Due to that we can replace output data which have bad analytical properties. The better analytical properties of modified data make possible, as we shall see below, to consider more constructive optimal problems than original ones.

Remark 2.3. The principal results of Theorem 2.2 remain valid if the conditions (2.18) and (2.19) are replaced by the conditions

$$(2.31)\qquad \bar{a}(0)=0,\ |\bar{a}(1)-a(1)|\le\delta_1,\ |\int_0^1 \bar{a}(s)ds - \int_0^1 a(s)ds|dt \le \delta_2.$$

A proof is similar to the proof in Remark 2.1.

Remark 2.4. For the classes $\mathbf{K}_n$, $n>2$, (see (1.7)) it is not difficult to obtain the analogous results. The following inequality

$$\frac{1}{2}\int_0^1 x_0^2(t)dt = \frac{1}{2}\int_0^1 (P_{n-1}(t) + \int_1^t (\int_1^{s_{n-1}} ...(\int_1^{s_1} x_0^{(n)}(s)ds)...)ds_{n-1})^2 dt$$

$$\ge \frac{1}{4}\int_0^1 P_{n-1}^2(t)dt - \frac{1}{2}\frac{M_n^2}{(2n+1)(n!)^2},$$

where

$$P_{n-1}(t) = x_0(1) + x_0'(1)(t-1) + ... + \frac{1}{(n-1)!}x_0^{(n-1)}(1)(t-1)^{n-1},$$

has an essential meaning for these classes (see the proof of the inequality (2.13) and the corresponding place in the proof of Theorem 2.2). Note that the integral

$$\int_0^1 P_{n-1}^2(t)dt = \int_0^1 (x_0(1) + x_0'(1)(t-1) + ... + \frac{1}{(n-1)!}x_0^{(n-1)}(1)(t-1)^{n-1})^2 dt$$

is a positive definite quadratic form with respect to $x_0(1)$,..., $x_0^{(n-1)}(1)$.

In conclusion let us remark that the requirements (2.20), (2.31) on $\bar{a}(t)$ for proximity $\bar{a}(t)$ to $a(t)$ in the case $n = 2$ are weaker than the requirements (2.14)-(2.15) in the case $n = 1$ and with growing n similar requirements are relaxed.

3. Reduction of maximum likelihood estimating to the problem of optimal road profile

Let us return to the problem of construction of maximum likelihood estimate $\hat{x}(t)$ in the class $\mathbf{K}_1$. This estimate can be found as a solution of the minimization problem (see (1.2))

$$I = \frac{1}{2}\int_0^1 x^2(t)dt - \int_0^1 x(t)da(t) \longrightarrow \min_{x(\cdot)\in\mathbf{K}_1} . \tag{3.1}$$

Consider also the following minimization problem

$$\bar{I} = \frac{1}{2}\int_0^1 x^2(t)dt - \int_0^1 x(t)d\bar{a}(t) \longrightarrow \min_{x(\cdot)\in\mathbf{K}_1} . \tag{3.2}$$

According to Theorem 2.1 or Remark 2.1 if $\bar{a}(\cdot)$ is close to $a(\cdot)$, then the solution $\bar{x}(t)$ of the problem (3.2) is close to the maximum likelihood estimate $\hat{x}(t)$. There are extensive possibilities for choice of the function $\bar{a}(t)$ such that the conditions (2.3)-(2.4) or (2.14)-(2.15) are satisfied. For instance, the function $\bar{a}(t)$ can easily be found as a piecewise linear function, which has a piecewise constant derivative.

Let $\bar{a}(t)$ in (3.2) satisfy (2.3)-(2.4) or (2.14)-(2.15) and be piecewise differentiable. Denote $\bar{a}'(t)$ by $b(t)$. Then the functional (3.2) transforms to the functional

$$\bar{I} = \frac{1}{2}\int_0^1 (x - b(t))^2 dt - \frac{1}{2}\int_0^1 b^2(t)dt,$$

and the following minimization problem appears (for the functional modified again we use without ambiguity the initial notation I)

$$I = \frac{1}{2}\int_0^1 (x - b(t))^2 dt \longrightarrow \min_{|x'|\leq M} . \tag{3.3}$$

The problem (3.3) is a problem of mean-square approximation by functions with bounded derivative. It can be interpreted as a problem of building road with profile $x(t)$ which cannot have steep ascents and descents and therefore $|x'(t)| \leq M$, $0 \leq t \leq 1$. The function

$b(t)$ is interpreted as a profile of a locality and the integral I as a cost of building. First this problem as the following problem of optimal control

$$I = \frac{1}{2}\int_0^1 (x - b(t))^2 dt \longrightarrow \min_{u:|u|\le M}, \tag{3.4}$$

$$x' = u \tag{3.5}$$

was studied by V.G. Boltyansky [2]. It has been studied in more detail and in more general form in the paper [3]. In particular in this paper the sufficiency of Pontryagin's maximum principle is proved when in place of one equation (3.5) one considers a general m-dimensional non autonomous linear system with r-dimensional control and instead of a functional with quadratic integrand one considers a functional with convex function. Besides in [3] the iterative procedure is recommended for finding optimal solution. Both V.G. Boltyansky and the authors of [3] made an assumption that $b(t)$ is piecewise differentiable. However this assumption is not essential; we are interested in the case where $b(t)$ is only piecewise continuous, since the simplest method of approximating $a(t)$ is realized by means of piecewise linear functions. As a result, as already mentioned, $b(t)$ will be piecewise constant. Therefore, but also for completeness of exposition we develop the required results from [3] with proofs, which are simplified substantially in the case considered.

Beforehand let us remark that the solution to problem (3.4)-(3.5) exists and is unique, which can be proved by traditional way in optimal control.

Let us write down necessary conditions for the optimal solution of problem (3.4)-(3.5). Pontryagin's function H has the form

$$H(t, x, u, p) = pu - \frac{\lambda_0}{2}(x - b(t))^2.$$

It is not difficult to prove that $\lambda_0 \neq 0$ and hence we can put $\lambda_0 = 1$. The optimal solution $u(t)$, $x(t)$ satisfies the system of differential equations

$$\frac{dx}{dt} = \frac{\partial H}{\partial p} = u, \tag{3.6}$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} = x - b(t), \tag{3.7}$$

the conditions of transversality

$$p(0) = 0,\ p(1) = 0, \tag{3.8}$$

and the maximum condition

$$p(t)u(t) = \max_{|v|\le M} p(t)v. \tag{3.9}$$

Theorem 3.1. *The solution* $u(t)$, $x(t)$ *of problem* (3.6)-(3.9) *is optimal for* (3.4)-(3.5). *Therefore, in view of the uniqueness of the optimal solution, the extreme solution is unique,* *in other words, the sufficiency of the maximum principle for problem* (3.4)-(3.5) *takes place.*

Proof. Let $\bar{u}(\cdot), \bar{x}(\cdot), \bar{p}(\cdot)$ be a solution of (3.6) - (3.9), i.e.

$$\frac{d\bar{x}}{dt} = \bar{u},\ \frac{d\bar{p}}{dt} = \bar{x} - b(t),\ \bar{p}(0) = \bar{p}(1) = 0,$$

$$\bar{p}(t)\bar{u}(t) = \max_{|v|\le M} \bar{p}(t)v.$$

Let $u(t)$ be an admissible control and $x(t)$ be some solution of equation (3.5). We have

$$I(x(\cdot)) = \frac{1}{2}\int_0^1 (x(t)-b(t))^2 dt = \int_0^1 (\frac{1}{2}(x(t)-\bar{x}(t))^2 - \bar{p}(t)u(t))dt$$
$$+ \int_0^1 (\frac{1}{2}(\bar{x}(t)-b(t))^2 - (\bar{x}(t)-b(t))\cdot \bar{x}(t))dt. \tag{3.10}$$

This equality follows from the obvious relation

$$\int_0^1 \bar{p}(t)u(t)dt = \int_0^1 \bar{p}(t)\frac{dx(t)}{dt}dt = -\int_0^1 x(t)(\bar{x}(t)-b(t))dt. \tag{3.11}$$

As the second integral on the right-hand side of (3.10) is a constant, the functional $I(x(\cdot))$ attains its minimum simultaneously with the functional

$$L(x(\cdot)) = \int_0^1 [\frac{1}{2}(x-\bar{x})^2 - \bar{p}(t)u]dt.$$

Since $\bar{p}(t)\bar{u}(t) \ge \bar{p}(t)u$ for arbitrary $|u| \le M$, the functional L obviously attains a minimum at $x(\cdot) = \bar{x}(\cdot),\ u(\cdot) = \bar{u}(\cdot)$. Theorem 3.1 is proved.

Let us adduce three lemmas which will be used in the next section.

Lemma 3.1. *Let the functions* $\tilde{x}(t), \tilde{p}(t)$ *satisfy* (3.7)-(3.8), *i.e.*

$$\frac{d\tilde{p}(t)}{dt} = \tilde{x}(t) - b(t),\ \tilde{p}(0) = \tilde{p}(1) = 0,$$

and the functions $u(t), x(t)$ *satisfy equation* (3.6), *i.e.*

$$\frac{dx(t)}{dt} = u(t).$$

Then we have (*analogously to* (3.11))

$$\int_0^1 \tilde{p}(t)u(t)dt = -\int_0^1 x(t)(\tilde{x}(t)-b(t))dt. \tag{3.12}$$

Proof. This assertion implies immediately by simple calculations.

Lemma 3.2. *To any admissible control* $u(t)$ *there corresponds a unique solution of the boundary value problem* (3.6)-(3.8).

Proof. Indeed

$$x(t) = x(0) + \int_0^t u(s)ds,$$

$$p(t) = \int_0^t [x(0) + \int_0^\tau u(s)ds - b(\tau)]d\tau.$$

From the condition $p(1) = 0$ we uniquely determine $x(0)$:

$$x(0) = -\int_0^1 \left[\int_0^\tau u(s)ds - b(\tau)\right] d\tau. \tag{3.13}$$

Lemma 3.3. *Let* $u(t)$ *and* $v(t)$ *be some admissible controls,* $u \neq v$, *and*

$$w(\alpha, t) = \alpha u(t) + (1-\alpha)v(t), \quad 0 \leq \alpha \leq 1.$$

Then there exist values $\overline{x}^0$, $\overline{\alpha}$, *which realize the minimal value of the function*

$$G(x^0, \alpha) = \frac{1}{2}\int_0^1 (x^0 + \int_0^t w(\alpha, \tau)d\tau - b(t))^2 dt$$

in the domain $-\infty < x_0 < +\infty$, $0 \leq \alpha \leq 1$. *The values* $\overline{x}^0$, $\overline{\alpha}$ *can be found by the following rule.*

First calculate the functions $\xi(t)$ *and* $\eta(t)$ *:*

$$\xi(t) = \int_0^t (u(s) - v(s))ds - \int_0^1 \int_0^\tau (u(s) - v(s))dsd\tau, \tag{3.14}$$

$$\eta(t) = \int_0^1 b(\tau)d\tau + \int_0^t v(s)ds - \int_0^1 \int_0^\tau v(s)dsd\tau. \tag{3.15}$$

Then find the constant β *:*

$$\beta = \frac{\int_0^1 \xi(t)(b(t) - \eta(t))dt}{\int_0^1 \xi^2(t)dt}.$$

Finally for $\bar{\alpha}$ *we have*

$$\overline{\alpha} = \begin{cases} \beta, & \text{if} \quad 0 < \beta < 1, \\ 0, & \text{if} \quad \beta \leq 0, \\ 1, & \text{if} \quad \beta \geq 1, \end{cases} \tag{3.16}$$

and for $\overline{x^0}$ *we have*

$$\overline{x}_0 = -\int_0^1 \left[\int_0^t w(\overline{\alpha}, s)ds - b(t)\right] dt. \tag{3.17}$$

Proof. The lemma concerns the minimization of the function $G(x_0, \alpha)$ which is quadratic in x^0 and α where x^0 varies from $-\infty$ to $+\infty$ and $0 \leq \alpha \leq 1$. Since the function G is unbounded for $x^0 \to \pm\infty$, the existence of values $\overline{x}_0, \overline{\alpha}$ easily follows. The function $G(x_0, \overline{\alpha})$ of one variable x_0 obviously takes a minimal value at $x_0 = \overline{x}_0$ which is calculated from the formula (3.17). Furthermore consider the function $G(x^0(\alpha), \alpha)$ depending on $0 \leq \alpha \leq 1$ where $x^0(\alpha)$ is determined by the right hand side of (3.17) with α substituted for $\overline{\alpha}$. Obviously $G(x^0(\alpha), \alpha)$ takes its minimal value at $\alpha = \overline{\alpha}$. This function is a quadratic polynomial in α which can immediately be calculated :

$$\begin{aligned} G(x^0(\alpha), \alpha) = {} & \frac{1}{2} \int_0^1 \xi^2(t) dt \cdot \alpha^2 \\ & - \int_0^1 \xi(t)(b(t) - \eta(t)) dt \cdot \alpha + \frac{1}{2} \int_0^1 (b(t) - \eta(t))^2 dt. \end{aligned}$$

From this we obtain the rule (3.16). Let us remark that in case $u \neq v$ the integral $\displaystyle\int_0^1 \xi^2(t) dt \neq 0$.

4. Iterative approximations

As a first approximation of the optimal control we take an arbitrary admissible control $u_1(t)$. The first approximation of the trajectory $x_1(t)$ and the function $p_1(t)$ are found according to Lemma 3.2. Let the k-th approximation be constructed: $u_k(t), x_k(t), p_k(t)$. Knowing $p_k(t)$, we find $v_k(t)$ from the condition

$$p_k(t) v_k(t) = \max_{|v| \leq M} p_k(t) v,$$

that is, in particular, one may put

$$v_k(t) = M \operatorname{sign} p_k(t) = \begin{cases} M, & p_k(t) > 0, \\ 0, & p_k(t) = 0, \\ -M, & p_k(t) < 0. \end{cases}$$

Then we apply Lemma 3.3 with

$$w(\alpha, t) = w_k(\alpha, t) = \alpha u_k(t) + (1 - \alpha) v_k(t),$$

assuming that $v_k \neq u_k$ (in the opposite case, as will be shown below, u_k is an optimal control). Let the point x^0_{k+1}, α_k realize the minimal value of the function

$$G_k(x^0, \alpha) = \frac{1}{2} \int_0^1 (x^0 + \int_0^t w_k(\alpha, \tau) d\tau - b(t))^2 dt.$$

Then the $(k+1)$-st approximation of the control is chosen in the form

$$u_{k+1}(t) = \alpha_k u_k(t) + (1 - \alpha_k) v_k(t), \tag{4.1}$$

and $x_{k+1}(t)$ and $p_{k+1}(t)$ are found according to Lemma 3.2. In this manner we construct a sequence $\{u_n(t), x_n(t), p_n(t)\}$. It is easy to see that $x_{k+1}(0) = x^0_{k+1}$ (by comparing the formulae (3.13) and (3.17)). Let us write down $x_{k+1}(t)$:

$$x_{k+1}(t) = -\int_0^1 (\alpha_k \int_0^t u_k(s)ds - \alpha_k b(t))dt \tag{4.2}$$

$$-\int_0^1 ((1-\alpha_k)\int_0^t v_k(s)ds - (1-\alpha_k)b(t))dt + \alpha_k \int_0^t u_k(s)ds$$

$$+(1-\alpha_k)\int_0^t v_k(s)ds = \alpha_k x_k(s) + (1-\alpha_k)\eta_k(s).$$

Let I_n be the value of the functional (3.3) at $x = x_n(t)$. By construction the sequence I_n is non increasing and is bounded from below by the least value of the functional I : $I_1 \geq I_2 \geq \cdots \geq I_n \geq \cdots \geq I_0$.

Theorem 4.1. *The sequence $x_n(t)$ converges uniformly on $[0,1]$ to the optimal trajectory.*

Proof. Let us first of all show that if $I_{n+1} = I_n$ then $x_n(t)$ is an optimal trajectory and consequently $x_{n+1}(t) = x_n(t)$. In view of the fact that $x_{k+1}(0) = x^0_{k+1}$ and $x_{k+1}(t)$ corresponds to the control $u_{k+1}(t)$ we have

$$I_{n+1} = G_n(x^0_{n+1}, \alpha_n) = \min G_n(x^0, \alpha).$$

Since $I_{n+1} = I_n$, the least value of the function G_n is attained at $\alpha_n = 1$ and $x^0 = x^0_n$. Since G_n is a quadratic parabola in α, it follows that

$$\frac{\partial G_n(x^0_n, 1)}{\partial \alpha} \leq 0,$$

i.e.,

$$\int_0^1 \left[\left(x^0_n + \int_0^t u_n(\tau)d\tau - b(t)\right)\int_0^t (u_n(\tau) - v_n(\tau))\, d\tau\right] dt \leq 0.$$

Putting in lemma 3.1

$$\widetilde{x}(t) - b(t) = x^0_n + \int_0^t u_n(\tau)d\tau - b(t),$$

$$u(t) = u_n(t) - v_n(t), \qquad x(t) = \int_0^t (u_n(\tau) - v_n(\tau))d\tau,$$

we obtain

$$\int_0^1 p_n(t)(v_n(t) - u_n(t))dt \leq 0.$$

But since $p_n(t)v_n(t) \geq p_n(t)u_n(t)$ for almost all t, the integral in the last relation is zero and we have almost everywhere

$$p_n(t)u_n(t) = p_n(t)v_n(t) = \max_{|u|\leq M} p_n(t)u.$$

Thus $u_n(t)$, $x_n(t)$, $p_n(t)$ satisfy the Pontryagin maximum principle (3.6)-(3.9) and according to Theorem 3.1 $x_n(t)$ is the optimal trajectory.

Consider now the general case, where for all n the strict inequality $I_{n+1} < I_n$ is fulfilled. In view of weak compactness of the unit ball in the space $L^2[0,1]$ there are sequences $u_{n_k}(t)$ and $v_{n_k}(t)$ which weakly converge to admissible controls $\overline{u}(t)$ and $\overline{v}(t)$. Since the sequence x_n^0 is bounded (this can easily be shown), we can assume without restricting generality that the sequences $x_{n_k}(t)$ and $p_{n_k}(t)$ converge uniformly on the interval $[0,1]$ to $\overline{x}(t)$ and $\overline{p}(t)$ respectively. Also without restricting generality we can assume that the sequence $x_{n_k+1}(t)$ is convergent ; its limit we denote by $\overline{\overline{x}}(t)$.

Let us demonstrate that $\overline{x}(t) = \overline{\overline{x}}(t)$. From the assumption $I_{n+1} < I_n$ it follows that $x_{n_k+1} \neq x_{n_k}$, hence $\alpha_{n_k} \neq 1$, i.e. $0 \leq \alpha_{n_k} < 1$. Since at $\alpha = \alpha_{n_k}$ the parabola $G_{n_k}(x^0_{n_k+1}, \alpha)$ attains its minimal value, we have

$$\frac{\partial G_{n_k}}{\partial \alpha}(x^0_{n_k+1}, \alpha_{n_k}) = 0$$

in the case $0 < \alpha_{n_k} < 1$ (the branches of the parabola point upwards), and

$$\frac{\partial G_{n_k}}{\partial \alpha}(x^0_{n_k+1}, 0) \geq 0$$

in the case $\alpha_{n_k} = 0$ (the branches point downwards). As a result we get

(4.3)
$$\frac{\partial G_{n_k}}{\partial \alpha}(x^0_{n_k+1}, \alpha_{n_k})$$
$$= \int_0^1 [(x^0_{n_k+1} + \int_0^t w_n(\alpha_{n_k}, \tau)d\tau - b(t)) \cdot \int_0^t (u_{n_k}(\tau) - v_{n_k}(\tau))d\tau]dt \geq 0.$$

Furthermore from (4.1)

$$u_{n_k} - v_{n_k} = \frac{1}{1-\alpha_{n_k}}(u_{n_k} - u_{n_k+1}),$$

and from (3.17) we obtain

(4.4)
$$\int_0^1 [x^0_{n_k+1} + \int_0^t w_n(\alpha_{n_k}, \tau)d\tau - b(t)]dt = 0.$$

Finally,

$$\int_0^t (u_{n_k}(\tau) - v_{n_k}(\tau))d\tau = \frac{1}{1-\alpha_{n_k}} \int_0^t (u_{n_k}(\tau) - u_{n_k+1}(\tau))d\tau$$
$$= \frac{1}{1-\alpha_{n_k}}(x_{n_k}(t) - x_{n_k+1}(t) - (x^0_{n_k} - x^0_{n_k+1})).$$

Inserting this integral into (4.3) and utilizing (4.4), we arrive at the inequality

(4.5)
$$\int_0^1 (x^0_{n_k+1} + \int_0^t w_n(\alpha_{n_k}, \tau)d\tau - b(t)) \cdot (x_{n_k}(t) - x_{n_k+1}(t))dt \geq 0.$$

Passing to the limit in (4.5), we get

$$\int_0^1 (\overline{\overline{x}}(t) - b(t))(\overline{x}(t) - \overline{\overline{x}}(t))dt \geq 0. \tag{4.6}$$

Furthermore

$$I_{n_k} - I_{n_k+1} = \frac{1}{2}\int_0^1 (x_{n_k}(t) - b(t))^2 dt - \frac{1}{2}\int_0^1 (x_{n_k+1}(t) - b(t))^2 dt$$

$$= \int_0^1 (x_{n_k+1}(t) - b(t)) \cdot (x_{n_k}(t) - x_{n_k+1}(t))dt + \frac{1}{2}\int_0^1 (x_{n_k}(t) - x_{n_k+1}(t))^2 dt.$$

Passing to the limit in this inequality, we obtain

$$\int_0^1 (\overline{\overline{x}}(t) - b(t))(\overline{x}(t) - \overline{\overline{x}}(t))dt + \frac{1}{2}\int_0^1 (\overline{x}(t) - \overline{\overline{x}}(t))^2 dt = 0. \tag{4.7}$$

In case $\overline{x}(t) \neq \overline{\overline{x}}(t)$ the relation (4.7) would imply the opposite inequality to (4.6). Hence we have in fact $\overline{\overline{x}}(t) = \overline{x}(t)$. This implies that $p_{n_k}(t) \to \bar{p}(t), \quad p_{n_k+1}(t) \to \bar{p}(t)$.

We note two further equalities. The first one is obtained from the relation (4.3) by means of Lemma 3.1:

$$\frac{\partial G_{n_k}}{\partial \alpha}(x^0_{n_k+1}, \alpha_{n_k}) = \int_0^1 p_{n_k+1}(t)(v_{n_k}(t) - u_{n_k}(t))dt. \tag{4.8}$$

We remind that

$$x_{n_k+1}(t) - b(t) = \frac{dp_{n_k+1}(t)}{dt}.$$

The second one is obtained by the some reasoning as with (4.5). At first we write

$$\frac{\partial G_{n_k}}{\partial \alpha}(x^0_{n_k+1}, \alpha_{n_k}) = \int_0^1 p_{n_k+1}(t)(v_{n_k}(t) - u_{n_k}(t))dt$$

$$= \frac{1}{1-\alpha_{n_k}}\int_0^1 (x_{n_k+1}(t) - b(t))(x_{n_k}(t) - x_{n_k+1}(t))dt$$

which implies the equality, required in the sequel:

$$\int_0^1 (x_{n_k+1}(t) - b(t))(x_{n_k}(t) - x_{n_k+1}(t))dt = (1-\alpha_{n_k})\int_0^1 p_{n_k+1}(t)(v_{n_k}(t) - u_{n_k}(t))dt. \tag{4.9}$$

The set of all indices n_k is such that either $\alpha_{n_k} = 0$ or $\alpha_{n_k} \neq 0$. If the set of indices n_k with $\alpha_{n_k} = 0$ is infinite, then passing to the limit in (4.9), we obtain

$$\int_0^1 \overline{p}(t)(\overline{v}(t) - \overline{u}(t))dt = 0. \tag{4.10}$$

If this set of indices is finite, then for the remaining indices the left part of (4.8) vanishes, and we again get (4.10).

The condition

$$p_{n_k}(t)v_{n_k}(t) = \max_{|v|\leq M} p_{n_k}(t)v$$

implies that for any admissible control $u(t)$ one has the inequality

$$p_{n_k}(t)v_{n_k}(t) \geq p_{n_k}(t)u(t).$$

Let E be an arbitrary measurable subset of the interval $[0,1]$. The following inequality holds:

$$\int_E p_{n_k}(t)v_{n_k}(t)dt \geq \int_E p_{n_k}(t)u(t)dt.$$

Passing to the limit here we get

$$\int_E \bar{p}(t)\overline{v}(t)dt \geq \int_E \overline{p}(t)u(t)dt.$$

Since E is arbitrary, this implies for almost all t

$$\bar{p}(t)\overline{v}(t) \geq \overline{p}(t)u(t). \tag{4.11}$$

In particular (4.11) holds also for $u(t) = \overline{u}(t)$. Therefore (4.10) implies that almost everywhere

$$\overline{p}(t)\overline{v}(t) = \overline{p}(t)\overline{u}(t).$$

Returning to (4.11), we obtain that

$$\overline{p}(t)\overline{u}(t) \geq \overline{p}(t)u(t)$$

for an arbitrary admissible control $u(t)$. This implies that $\overline{u}(t), \overline{x}(t), \overline{p}(t)$ satisfy the necessary conditions of the maximum principle. As a consequence of Theorem 3.1 $\overline{x}(t)$ is the optimal trajectory. In view of the uniqueness the sequence $x_n(t)$ itself converges to the optimal trajectory uniformly. Theorem 4.1 is proved.

Remark 4.1. The sequence $u_n(t)$ can be seen to converge to the optimal control weakly.

Remark 4.2. Since obviously $u_k(t)$ is a piecewise constant function, $x_k(t)$ is always a piecewise linear continuous function. Therefore if $b(t)$ is a piecewise constant or continuous piecewise linear function, $p_k(t)$ is a quadratic spline (of defect 2 or 1). The knots of this spline are the switching points of $u_k(t)$ and the non regular points of the function $b(t)$.

5. Inserting a parameter

In this section we give another approach for constructive solving the problem (3.4)-(3.5). Let us consider the following problem of optimal control

$$I = \frac{1}{2}\int_0^1 ((x-b(t))^2 + \lambda u^2)dt \to \min_{|u|\leq M}, \tag{5.1}$$

$$\frac{dx}{dt} = u, \tag{5.2}$$

which depends on the parameter $\lambda \geq 0$. For $\lambda = 0$ the problem coincides with (3.4)-(3.5). Clearly, the solution of problem (5.1)-(5.2) for small positive λ is close to the required solution of (3.4)-(3.5).

Pontryagin's function H of the problem (3.4)-(3.5) has the form

$$H(t, x, u, p) = pu - \frac{1}{2}(x - b(t))^2 - \frac{1}{2}\lambda u^2.$$

Necessary conditions (it can be proved that they are sufficient as well) for the optimal solution under $\lambda > 0$:

$$\frac{dx}{dt} = \frac{\partial H}{\partial p} = u, \tag{5.3}$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} = x - b(t),\ p(0) = p(1) = 0, \tag{5.4}$$

$$p(t)u(t) - \frac{1}{2}\lambda u^2(t) = \max_{|v| \leq M}(p(t)v - \frac{1}{2}\lambda v^2). \tag{5.5}$$

The condition (5.5) gives for u the following expression:

$$u = u(p; \lambda) := \begin{cases} -M,\ p < -\lambda M, \\ p/\lambda\ ,\ |p| \leq \lambda M, \\ M,\ p > \lambda M. \end{cases} \tag{5.6}$$

Therefore to find the optimal solution we have to solve the boundary value problem

$$\frac{dx}{dt} = u(p; \lambda),\ \frac{dp}{dt} = x - b(t),\ p(0) = p(1) = 0. \tag{5.7}$$

A little below we justify that for all sufficiently large λ the restriction $|p| \leq \lambda M$ is fulfilled and consequently problem (5.7) acquires the form

$$\frac{dx}{dt} = \frac{p}{\lambda},\ \frac{dp}{dt} = x - b(t),\ p(0) = p(1) = 0. \tag{5.8}$$

Problem (5.8) has the following explicit solution

$$x = \frac{1}{\sqrt{\lambda}}(\sinh\frac{1}{\sqrt{\lambda}})^{-1} \cdot \cosh\frac{t}{\sqrt{\lambda}}\int_0^1 \cosh\frac{1-\tau}{\sqrt{\lambda}}b(\tau)d\tau - \frac{1}{\sqrt{\lambda}}\int_0^t \sinh\frac{t-\tau}{\sqrt{\lambda}}b(\tau)d\tau,$$

$$p = (\sinh\frac{1}{\sqrt{\lambda}})^{-1} \cdot \sinh\frac{t}{\sqrt{\lambda}}\int_0^1 \cosh\frac{1-\tau}{\sqrt{\lambda}}b(\tau)d\tau - \int_0^t \cosh\frac{t-\tau}{\sqrt{\lambda}}b(\tau)d\tau.$$

Now it can be verified that the above-mentioned restriction is fulfilled in fact if, for example,

$$\sqrt{\lambda}\exp(-\frac{1}{\sqrt{\lambda}}) > \frac{B}{M}\ ,$$

where $B := \max_{0 \leq s \leq 1}|b(s)|$.

Let us denote $x(t;\lambda,x_0)$, $p(t;\lambda,x_0)$ the solution to the Cauchy problem

$$\frac{dx}{dt}=u(p;\lambda),\ \frac{dp}{dt}=x-b(t),\ x(0)=x_0,\ p(0)=0. \tag{5.9}$$

To solve the boundary value problem (5.7) for a $\lambda>0$ it is necessary to find $x_0(\lambda)$ such that

$$p(1;\lambda,x_0(\lambda))=0. \tag{5.10}$$

The transcendental equation (5.10) can be solved easily (for example, by the chord method) if an initial approximation for $x_0(\lambda)$ is known accurately enough.

Let $x_0(\lambda)$ be known. Then $x_0(\lambda-\Delta\lambda)$ can be found from the equation

$$p(1;\lambda-\Delta\lambda,x_0(\lambda-\Delta\lambda))=0$$

if we take $x_0(\lambda)$ as an initial approximation for $x_0(\lambda-\Delta\lambda)$. Thus, knowing the optimal solution for some $\lambda>0$ (fortunately, we do know it for a large $\lambda>0$), we can find it for $\lambda-\Delta\lambda$ in a constructive manner. Resting on these ideas, it is not difficult to construct a numerical procedure for solving the problem (5.1)-(5.2) for any $\lambda>0$ and consequently for approximate solving the required problem (3.4)-(3.5).

6. Some generalizations

The problem of finding a maximum likelihood estimate $\hat{x}(t)$ in each class $\mathbf{K}_n$ (see (1.7)), $n\geq 2$, is solved analogously. After substituting $a(t)$ by a nearby $\overline{a}(t)$ such that there exists the piecewise continuous derivative $\overline{a}'(t)=b(t)$, this problem is also reduced to the "problem of finding the optimal road profile". For example, in the case $n=2$ we obtain the problem of minimization of the functional

$$I=\frac{1}{2}\int_0^1(x_1-b(t))^2dt\to\min_{|u|\leq M} \tag{6.1}$$

$$\frac{dx_1}{dt}=x_2, \tag{6.2}$$

$$\frac{dx_2}{dt}=u.$$

The problem (6.1)-(6.2) can be solved with using [3] as the problem (3.4)-(3.5) was done above. And due to Theorem 2.2 the optimal solution $\bar{u}(t)$, $\bar{x}_1(t)$, $\bar{x}_2(t)$ of the problem is such that $\bar{x}_1(t)$ is close to $\hat{x}(t)$.

The same approach is possible also in the case of stronger information on the unknown signal. For instance, it may be known that the signal is a non decreasing function with the first derivative bounded from above. Then

$$\mathbf{K}_1^*=\Big\{x(\cdot):x(t)\text{ is absolutely continuous and }0\leq x'(t)\leq M\Big\},$$

and the optimal control problem takes the form

$$I = \frac{1}{2}\int_0^1 (x - b(t))^2 dt \to \min_{0 \le u \le M} \ , \tag{6.3}$$

$$x' = u. \tag{6.4}$$

Let us introduce a new control v and a new phase variable y:

$$v = u - \frac{M}{2}, \qquad y = x - \frac{M}{2}t.$$

Then the problem (6.3)-(6.4) transforms to

$$I = \frac{1}{2}\int_0^1 (y - c(t))^2 dt \to \min_{|v| \le M/2} \ ,$$

$$y' = v,$$

where $c(t) = b(t) - Mt/2$, which coincides with the problem (3.4)-(3.5).

Now consider the class

$$\mathbf{K}_2^* = \Big\{ x(\cdot) : \ x'(t) \text{ is absolutely continuous and } \quad 0 \le x''(t) \le M \Big\}$$

which corresponds to information on the signal being a convex function with bounded second derivative. As above it can be reduced to the problem

$$I = \frac{1}{2}\int_0^1 (y_1 - c(t))^2 dt \to \min_{|v| \le M/2} \ ,$$

$$\frac{dy_1}{dt} = y_2,$$

$$\frac{dy_2}{dt} = v,$$

where

$$v = u - \frac{M}{2}, \ y_1 = x - \frac{Mt^2}{4}, \ c(t) = b(t) - \frac{Mt^2}{4},$$

which coincides with the problem (6.1)-(6.2).

Analogously one treats the case where it is known that there exists absolutely continuous $x^{(n-1)}(t)$, and $\quad 0 \le x^{(n)}(t) \le M$. Such a class appears if it is known that the signal does not have more than n pieces of monotonicity (and, of course, if it is sufficiently smooth and its n-th derivative is subject to the bounds indicated).

Another quite natural information on the signal would be

$$\mathbf{K} = \Big\{ x(\cdot) : A \le x(t) \le B, \ \ x^{(n-1)}(t) \text{ is absolutely continuous and } \quad 0 \le x^{(n)}(t) \le M \Big\},$$

i.e. besides the fact that the signal does not have more than n pieces of monotonicity it is known that it is in a certain band. This problem can also be reduced to a typical optimal control problem but already with bounded phase variables. To find a sufficiently constructive solution of such problems is a more complicated task.

7. Maximum likelihood estimate for signal of Sobolev type

The optimal control problem (1.4)-(1.5) in the class (1.6) has the form

$$I = \int_0^1 (\frac{1}{2}x^2 + a(t)u)dt - a(1)x(1) \longrightarrow \min, \tag{7.1}$$

$$x' = u, \tag{7.2}$$

$$\frac{1}{2}\int_0^1 (\alpha x^2 + u^2)dt - M \leq 0. \tag{7.3}$$

It is not difficult to prove the existence of a solution to this problem for an arbitrary continuous function $a(t)$ and constants $\alpha \geq 0$, $M > 0$.

Let us write down necessary conditions for an optimal solution of the problem (7.1)-(7.3) (we use the book [1] in this connection). There exist nonnegative constants $\lambda_0 \geq 0$, $\lambda_1 \geq 0$ and a function $p(t)$, $0 \leq t \leq 1$, which cannot vanish simultaneously such that Pontryagin's function

$$H(t,x,u,p) = pu - \lambda_0(\frac{1}{2}x^2 + a(t)u) - \lambda_1(\frac{1}{2}\alpha x^2 + \frac{1}{2}u^2)$$

receives the maximal value under optimal control, i.e. the equality

$$p - \lambda_0 a(t) - \lambda_1 u = 0 \tag{7.4}$$

is fulfilled.

The optimal solution $u(t)$, $x(t)$ satisfies the system of differential equations

$$x' = \frac{\partial H}{\partial p} = u \tag{7.5}$$

$$p' = -\frac{\partial H}{\partial x} = \lambda_0 x + \alpha\lambda_1 x. \tag{7.6}$$

In addition the conditions of transversality

$$p(0) = 0,\ p(1) = \lambda_0 a(1) \tag{7.7}$$

and the condition of the complementary slackness

$$\lambda_1 \cdot (\frac{1}{2}\int_0^1 (\alpha x^2 + u^2)dt - M) = 0 \tag{7.8}$$

are fulfilled.

Using the necessary conditions (7.4)-(7.8), we can find the optimal solution of the problem (7.1)-(7.3). Let us prove first that

$$\lambda_0 > 0.$$

Indeed, if $\lambda_0 = 0$ then $\lambda_1 \neq 0$ since otherwise from (7.6)-(7.7) we have $p \equiv 0$, which is impossible since λ_0, λ_1, p cannot vanish simultaneously. So, if $\lambda_0 = 0$, then $\lambda_1 > 0$. Therefore (7.4) implies $u = p/\lambda_1$, and from (7.7) $p(0) = p(1) = 0$. The system (7.5)-(7.6) gives $p'' = \alpha p$, $p(0) = p(1) = 0$. Since $\alpha \geq 0$, we have $p \equiv 0$ and then $u \equiv 0$, $\alpha x \equiv 0$.

Consequently both multipliers in (7.8) are nonzero (remember that $M > 0$) and the condition (7.8) is violated. So $\lambda_0 > 0$ and we can put

$$\lambda_0 = 1\,.$$

Further, according to $\lambda_1 \geq 0$ we have two cases: $\lambda_1 = 0$ or $\lambda_1 > 0$. The case $\lambda_1 = 0$ yields $p = a$ in view of (7.4). Hence, from (7.6) $x = a'$. Such a solution is possible only if there exists $a'' \in \mathbf{L}_2[0,1]$ and

$$\frac{1}{2}\int_0^1 (\alpha a'^2(t) + a''^2(t))dt \leq M.$$

This is interesting in itself. But for the problems considered here this case must be excluded beforehand since (1.1) implies that $a(\cdot)$ is a non-differentiable function.

Thus we have $\lambda_1 > 0$. Consequently (we write λ for λ_1)

$$p'' = \frac{1+\alpha\lambda}{\lambda}\cdot(p-a), \quad p(0) = 0,\ p(1) = a(1), \tag{7.9}$$

$$\frac{1}{2}\int_0^1 (\alpha x^2(t) + u^2(t))dt = M, \tag{7.10}$$

$$u = \frac{p-a}{\lambda}, \qquad x = \frac{1}{1+\alpha\lambda}\,p'\,. \tag{7.11}$$

The solution to the boundary value problem (7.9) is of form

$$p(t) = C\sinh\beta t + \beta\int_0^t a(s)\sinh\beta(s-t)ds, \tag{7.12}$$

where

$$\beta = \left(\frac{1+\alpha\lambda}{\lambda}\right)^{1/2}, \quad C = \frac{1}{\sinh\beta}\cdot(a(1) + \beta\int_0^1 a(s)\,\sinh\beta(1-s)ds). \tag{7.13}$$

Let us prove now that the unknown constant λ can be found uniquely from (7.10) where u and x are from (7.11)-(7.13). Thereby it will be proved the uniqueness of the extreme solution for the optimal problem (7.1)-(7.3) or, that is the same, the sufficiency of the necessary conditions (7.4)-(7.8). To this end consider the following problem

$$J_\mu = \int_0^1 (\frac{1}{2}x^2 + a(t)u)dt + \frac{\mu}{2}\int_0^1 (\alpha x^2 + u^2)dt - a(1)x(1) \longrightarrow \min, \tag{7.14}$$

$$x' = u, \tag{7.15}$$

which is a problem without restrictions on control.

It is not difficult to obtain that for every $\mu > 0$ the problem (7.14)-(7.15) has a unique optimal solution $u(t)$, $x(t)$ which has a form like (2.11)

$$u = \frac{p-a}{\mu}, \qquad x = \frac{1}{1+\alpha\mu}\,p'\,. \tag{7.16}$$

The function $p(t)$ in (7.16) has the form (7.12)-(7.13) with

$$\beta = \left(\frac{1+\alpha\mu}{\mu}\right)^{1/2}. \tag{7.17}$$

Introduce the functional

$$L = \frac{1}{2}\int_0^1 (\alpha x^2 + u^2)dt. \tag{7.18}$$

If we prove that the functional (7.18) calculated along optimal solution (7.16)-(7.17) strongly monotonically decreases as a function of μ, then the univalent solvability of the equation (7.10) with respect to λ will be proved.

Denote the functionals I, L, and J_ν calculated along the optimal solution $u_\mu(\cdot)$, $x_\mu(\cdot)$ of problem (7.14)-(7.15) by

$$I(\mu) = I(u_\mu(\cdot), x_\mu(\cdot)),\ L(\mu) = L(u_\mu(\cdot), x_\mu(\cdot)),\ J_\nu(\mu) = J_\nu(u_\mu(\cdot), x_\mu(\cdot)) = I(\mu) + \nu L(\mu).$$

Let $0 < \mu_1 < \mu_2$. Due to uniqueness of optimal solution of problem (7.14)-(7.15), we have

$$J_{\mu_2}(\mu_2) = J_{\mu_2}(u_{\mu_2}(\cdot), x_{\mu_2}(\cdot)) < J_{\mu_2}(u_{\mu_1}(\cdot), x_{\mu_1}(\cdot)) = J_{\mu_2}(\mu_1),$$
$$J_{\mu_1}(\mu_1) = J_{\mu_1}(u_{\mu_1}(\cdot), x_{\mu_1}(\cdot)) < J_{\mu_1}(u_{\mu_2}(\cdot), x_{\mu_2}(\cdot)) = J_{\mu_1}(\mu_2)$$

Consequently

$$\begin{aligned} J_{\mu_2}(\mu_2) &= I(\mu_2) + \mu_2 L(\mu_2) \\ &< I(u_{\mu_1}(\cdot), x_{\mu_1}(\cdot)) + \mu_2 L(u_{\mu_1}(\cdot), x_{\mu_1}(\cdot)) = I(\mu_1) + \mu_2 L(\mu_1). \end{aligned} \tag{7.19}$$

Analogously

$$I(\mu_1) + \mu_1 L(\mu_1) < I(\mu_2) + \mu_1 L(\mu_2). \tag{7.20}$$

From (7.19) and (7.20) the inequality

$$\mu_1(L(\mu_1) - L(\mu_2)) < I(\mu_2) - I(\mu_1) < \mu_2(L(\mu_1) - L(\mu_2)) \tag{7.21}$$

follows.

But (7.21) is possible if and only if

$$L(\mu_1) > L(\mu_2).$$

The strong monotonicity of the function $L(\mu)$ is proved.

As a result we obtain the following theorem.

Theorem 7.1. *The maximum likelihood estimate* $\hat{x}(\cdot)$ *in the model* (1.1) *in the class of signals* (1.6) *is given by formula*

$$\hat{x}(t) = \frac{1}{1+\alpha\lambda} p'(t) \tag{7.22}$$

where $p(t)$ *and* λ *are found uniquely from* (7.10)-(7.13).

Remark 7.1. The estimate (7.22) is nonlinear with respect to observation $a(\cdot)$ since λ, which must be found from (7.10)-(7.13), depends on $a(\cdot)$. At the same time the estimate (see (7.16)-(7.17))

$$\hat{x}(t) = \frac{1}{1+\alpha\mu} p'(t) \tag{7.23}$$

for every fixed μ is linear with respect to observation $a(\cdot)$. The estimate (7.23) can be treated due to the problem (7.14)-(7.15) as maximum likelihood estimate with penalty.

Remark 7.2. It is possible to consider analogously the problem (1.2) in other Sobolev's classes of functions, for instance, in the class

$$\mathbf{K} = \left\{ x(\cdot) : \exists x''(\cdot) \in \mathbf{L}_2[0,1] \, , \; \frac{1}{2} \int_0^1 (\alpha_0 x^2(t) + \alpha_1 x'^2(t) + x''^2(t)) dt \leq M \right\},$$

where $\alpha_0 \geq 0$, $\alpha_1 \geq 0$, $M > 0$.

References

[1] Alexeev V.M., Tichomirov V.M., Fomin S.V. Optimal Control. Moscow, Nauka, 1979.

[2] Boltyanskii V.G. Application of the theory of optimal processes to problems of approximation of functions. Trudy of Steclov Math. Inst., v.60, 1961, pp. 82-95.

[3] Gitenjova Yu.A., Milstein G.N. Minimization of convex functional by free trajectories of linear system. Differential equations, v.3, no.12, 1967, pp. 2081-2093.

[4] Ibragimov I.A., Khasminsky R.Z. Asymptotic Theory of Estimation. Moscow, Nauka, 1979.

[5] Nemirovsky A.S., Polyak B.T., Tsibakov A.B. Signal processing via the nonparametric maximum likelihood method. Problemy Peredachi Informacii, 20(1984), pp. 29-46.

[6] Nemirovski A.S., Polyak B.T., Tsybakov A.B. (1985). Rate of convergence of nonparametric estimates of maximum likelihood type. Problemy Peredachi Informacii, 21(1985), pp. 258-272.

[7] Tsirel'son B.S. A geometrical approach to maximum likelihood estimation for infinite-dimensional Gaussian location. I; II; III, Theory Probab. Appl., 27(1982), pp. 411-418; 30(1985), pp. 820-828; 31(1986), pp. 470-483.