Siberian Mathematical Journal, Vol. 59, *No.* 1, *pp.* 11–21, 2018

RECOVERING LINEAR OPERATORS AND LAGRANGE FUNCTION MINIMALITY CONDITION

© **A. V. Arutyunov and K. Yu. Osipenko** UDC 517.984

Abstract: This article concerns the recovery of the operators by noisy information in the case that their norms are defined by integrals over infinite intervals. We study the conditions under which the dual extremal problem (often nonconvex) can be solved using the Lagrange function minimality condition.

DOI: 10.1134/S0037446618010020

Keywords: optimal recovery, linear operator, extremal problem, Lagrange function

Introduction

Take a linear space X, normed linear spaces $Y_0, Y_1, \dots, Y_m$, and linear operators $I_j : X \to Y_j$ for $j = 0, 1, \dots, m$. Fix an integer k with $0 \le k < m$ and reals $\delta_j \ge 0$ for $j = 1, \dots, m$. Consider the optimal recovery problem for the operator $I_0 : X \to Y_0$ on the set

$$W = \{x \in X : \|I_j x\|_{Y_j} \le \delta_j,\ j = 1, \dots, k\} \tag{1}$$

from the values of $I_{k+1}, \dots, I_m$ known with some error; for $k = 0$ we put $W = X$. Assume that for each $x \in W$ we know some vector $y = (y_{k+1}, \dots, y_m) \in Y_{k+1} \times \dots \times Y_m$ such that $\|I_j x - y_j\|_{Y_j} \le \delta_j$ for $j = k+1, \dots, m$. Given y, we seek the element of Y_0 closest in the metric of this space to $I_0 x$.

Let us proceed to a more precise statement of the problem. Each method that for a given vector y indicates an approximation to the element $I_0 x$ amounts to a mapping from $Y_{k+1} \times \dots \times Y_m$ into Y_0. We consider all possible methods or, in other words, all possible mappings $\varphi : Y_{k+1} \times \dots \times Y_m \to Y_0$. For each mapping φ of this sort define its recovery error as

$$e(I, \delta, \varphi) = \sup_{\substack{x \in W,\ y \in Y_{k+1} \times \dots \times Y_m \\ \|I_j x - y_j\|_{Y_j} \le \delta_j,\ j = k+1, \dots, m}} \|I_0 x - \varphi(y)\|_{Y_0},$$

where $I = (I_0, I_1 \dots, I_m)$ and $\delta = (\delta_1, \dots, \delta_m)$. We have to find the error of optimal recovery defined as

$$E(I, \delta) = \inf_{\varphi : Y_{k+1} \times \dots \times Y_m \to Y_0} e(I, \delta, \varphi), \tag{2}$$

as well as the methods, if existent, at which this infimum is attained; these methods are called *optimal*.

Actually, instead the operator I_0 itself, which is given, we recover its values at the elements of W from noisy information about them. But this problem is traditionally called the *recovery problem for the operator* I_0.

In the simplest case, when $I_0, I_{k+1}, \dots, I_m$ are linear functionals, while W, in contrast to (1), is an arbitrary set in X and $\delta_{k+1} = \dots = \delta_m = 0$, this problem was posed by Smolyak [1]. He proved that

The authors were supported by the Ministry for Education and Science of Russia (Project 1.962.2017/4.6), the Peoples' Friendship University of Russia (Program 5–100), and the Russian Foundation for Basic Research (Projects 17–51–52022; 18–01–00106; 17–51–12064). Lemmas 1 and 2 and Theorems 1 and 2 were obtained by the first author with the support of the Russian Science Foundation (Project 17–11–01168).

Moscow. Translated from *Sibirskiĭ Matematicheskiĭ Zhurnal*, Vol. 59, No. 1, pp. 15–28, January–February, 2018; DOI: 10.17377/smzh.2018.59.102. Original article submitted March 12, 2017.

for every centrally symmetric convex set W there exists a linear optimal recovery method. Many articles generalize this statement; see [2–8], as well as the references therein.

In [9], a general result pertaining to the existence of a linear optimal method in the case that $m = 2$, while Y_0, Y_1, and Y_2 are Hilbert spaces, was justified, and first concrete results on the recovery of linear operators were obtained. This topic was further advanced in [10–12].

Basing on the second-order necessary conditions for extremum for abnormal problems, [13, 14] developed a method, whose use in this article enables us to obtain a series of results on the optimal recovery of linear operators.

1. The Dual Problem and Lagrange Function Minimality Condition

Refer as the *dual problem* to (2) to the extremal problem

$$\|I_0 x\|_{Y_0} \to \max, \quad \|I_j x\|_{Y_j} \le \delta_j,\ j = 1, \dots, m, \quad x \in X. \tag{3}$$

The objective value of this problem yields a lower bound for $E(I, \delta)$ due to the following well-known proposition; see [10, Lemma 1] for instance.

Lemma 1. *We have*

$$E(I, \delta) \ge \sup_{\substack{x \in X \\ \|I_j x\|_{Y_j} \le \delta_j,\ j=1,\dots,m}} \|I_0 x\|_{Y_0}.$$

Assume that $Y_0, Y_1, \dots, Y_m$ are Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_{Y_j}$ for $j = 0, 1, \dots, m$. Then it is convenient to pass to the squared objective value of (3) and consider the problem

$$\|I_0 x\|_{Y_0}^2 \to \max, \quad \|I_j x\|_{Y_j}^2 \le \delta_j^2, \quad x \in X. \tag{4}$$

Problem (4) is equivalent to the following:

$$q_0(x) \to \min, \quad q_j(x) \le \delta_j^2,\ j = 1, \dots, m, \quad x \in X, \tag{5}$$

where the quadratic forms q_j are defined as

$$q_0(x) = -\langle I_0 x, I_0 x \rangle_{Y_0}, \quad q_j(x) = \langle I_j x, I_j x \rangle_{Y_j},\ j = 1, \dots, m. \tag{6}$$

We are interested in which cases in the quadratic problem (5) the Lagrange function minimality condition is satisfied which is understood in the following strong sense.

Given real functions $f_j : X \to \mathbb{R}$, say that in the extremal problem

$$f_0(x) \to \min, \quad f_j(x) \le 0,\ j = 1, \dots, m, \tag{7}$$

the Lagrange function minimality condition is satisfied whenever there exist Lagrange multipliers $\lambda^j \ge 0$ for which

$$\inf_{x \in X} L(x, \lambda) = \inf_{\substack{x \in X \\ f_j(x) \le 0,\ j=1,\dots,m}} f_0(x).$$

Here L is the Lagrange function defined as

$$L(x, \lambda) = f_0(x) + \sum_{j=1}^{m} \lambda^j f_j(x), \quad \lambda = (\lambda^1, \dots, \lambda^m).$$

The Lagrange function minimality condition need not always hold. We study (5), which is a particular case of the general problem (7), but even in it the condition can be violated. Problem (5) yields the corresponding example for $m = 3$, $X = \mathbb{R}^2$, $x = (x_1, x_2) \in \mathbb{R}^2$, $\delta_j = 1$, and

$$q_0(x) = -\left(x_1^2 + x_2^2\right), \quad q_1(x) = (x_1 + 2x_2)^2, \quad q_2(x) = (x_1 - 2x_2)^2, \quad q_3(x) = 9x_1^2.$$

Consider (5) with the quadratic forms q_j of a more general form than those considered above. Namely, assume that q_j are of the form

$$q_j(x) = \langle Q_j x, x \rangle, \quad j = 0, \dots, m, \tag{8}$$

where $Q_j : X \to X^*$ are given linear operators.

Proposition 1. *Suppose that the infimum in (5) is finite and the Lagrange function minimality condition holds, i.e., there exist $\lambda^j \geq 0$, such that*

$$\inf_{x\in X} L(x,\lambda) = \inf_{\substack{x\in X\\ q_j(x)\leq\delta_j^2,\ j=1,\dots,m}} q_0(x), \quad \lambda = (\lambda^1,\dots,\lambda^m), \tag{9}$$

where the Lagrange function L is defined as

$$L(x,\lambda) = q_0(x) + \sum_{j=1}^m \lambda^j \big(q_j(x) - \delta_j^2\big).$$

Then

$$\inf_{\substack{x\in X\\ q_j(x)\leq\delta_j^2,\ j=1,\dots,m}} q_0(x) = -\sum_{j=1}^m \lambda^j\delta_j^2 = -\min_{\mu\in\Lambda}\sum_{j=1}^m \mu^j\delta_j^2, \quad \mu = (\mu^1,\dots,\mu^m), \tag{10}$$

where Λ consists of those Lagrange multipliers μ for which $\mu^j \geq 0$ for $j = 1,\dots,m$ and the quadratic form $q_0 + \mu^1 q_1 + \dots + \mu^m q_m$ is nonnegative definite. Moreover,

$$\inf_{\substack{x\in X\\ q_j(x)\leq\delta_j^2,\ j=1,\dots,m}} q_0(x) = \inf_{\substack{x\in X\\ \sum_{j=1}^m \lambda^j q_j(x)\leq\sum_{j=1}^m \lambda^j\delta_j^2}} q_0(x). \tag{11}$$

Proof. Since the infimum in (5) is finite, it follows that

$$q_0(x) + \sum_{j=1}^m \lambda^j q_j(x) \geq 0 \tag{12}$$

for all $x \in X$. Consequently,

$$\inf_{x\in X} L(x,\lambda) = -\sum_{j=1}^m \lambda^j\delta_j^2.$$

Take $\mu \in \Lambda$ and an admissible $x \in X$ in (5). Then

$$q_0(x) \geq L(x,\mu) \geq -\sum_{j=1}^m \mu^j\delta_j^2.$$

Taking the lower bound over all admissible elements, we obtain

$$-\sum_{j=1}^m \lambda^j\delta_j^2 = \inf_{\substack{x\in X\\ q_j(x)\leq\delta_j^2,\ j=1,\dots,m}} q_0(x) \geq -\sum_{j=1}^m \mu^j\delta_j^2.$$

This implies the second equality in (10).

Let us establish (11). Suppose that $x \in X$ and

$$\sum_{j=1}^m \lambda^j q_j(x) \leq \sum_{j=1}^m \lambda^j\delta_j^2.$$

Then (12) yields

$$q_0(x) \geq q_0(x) + \sum_{j=1}^m \lambda^j q_j(x) - \sum_{j=1}^m \lambda^j\delta_j^2 \geq -\sum_{j=1}^m \lambda^j\delta_j^2.$$

Consequently,

$$\inf_{\substack{x\in X\\ \sum_{j=1}^m \lambda^j q_j(x)\leq\sum_{j=1}^m \lambda^j\delta_j^2}} q_0(x) \geq -\sum_{j=1}^m \lambda^j\delta_j^2.$$

On the other hand,

$$\inf_{\substack{x\in X\\ \sum_{j=1}^m \lambda^j q_j(x)\le \sum_{j=1}^m \lambda^j\delta_j^2}} q_0(x) \le \inf_{\substack{x\in X\\ q_j(x)\le\delta_j^2,\ j=1,\dots,m}} q_0(x) = -\sum_{j=1}^m \lambda^j\delta_j^2. \quad \square$$

On assuming that X is a Hilbert space, we present a condition that ensures the fulfillment of the above minimality conditions for the Lagrange function for the quadratic problem (5), in which the quadratic forms q_j look like (8), where $Q_j : X \to X$ are given symmetric linear operators. Observe that the quadratic forms q_j are continuous because the symmetric operators Q_j are continuous by the Hellinger–Toeplitz Theorem.

Put

$$\overline{\Lambda} = \left\{\bar\lambda = (\lambda^0,\dots,\lambda^m) : \lambda^j \ge 0,\ j = 0,\dots,m,\quad \sum_{j=0}^m \lambda^j = 1\right\}.$$

Recall that the index of a quadratic form, denoted by ind, is the maximal dimension of the linear subspace on which this form is negative definite. This index can also take the infinite value.

Following [14], say that a system of quadratic forms q_j for $j = 0, 1, \dots, m$, *satisfies condition* $\mathscr{A}$ whenever for all $\bar\lambda \in \overline{\Lambda}$ the quadratic form defined by the relation $\lambda^0 q_0(x) + \dots + \lambda^m q_m(x)$, $x \in X$, is either nonnegative definite or has index greater than m.

Theorem 1. *If X is a Hilbert space, the infimum in* (5) *is finite, and condition $\mathscr{A}$ holds, then the Lagrange function minimality condition is fulfilled for this problem.*

Proof. We will follow [14]. Put

$$D = \{x \in X : q_j(x) \le \delta_j^2,\ j = 1,\dots,m\},\quad \kappa = \inf_{x\in D} q_0(x).$$

The nonempty set D is closed since q_j is continuous. Hence, D, with the metric induced from the complete space X, is itself a complete metric space. Therefore, we can apply to (5) the smooth variational principle of Ioffe and Tikhomirov, Theorem 1 of [15]; see also Theorem 2.6.5 in [16]. Take some $\varepsilon > 0$. Use Theorem 1 of [15], putting

$$\lambda = \sqrt[3]{\varepsilon},\quad \alpha_n = \sqrt[3]{\varepsilon}2^{-(n+1)},\quad \beta_n = 2^{-(3n+2)},\quad \varphi_{x,\alpha}(\xi) = 1 - \left|\frac{\xi - x}{\alpha}\right|^2.$$

Take some $w \in D$ with $q_0(w) \le \kappa + \varepsilon$. By the same theorem there exist a nonnegative sequence $\{\theta_n\}$ and a sequence $\{x_n\} \subset D$, depending on ε, converging to some point $x_* \in D$, such that

$$\theta_n \le 2^{-n},\quad |x_n - w| \le \sqrt[3]{\varepsilon},\quad q_0(x_n) \le q_0(w), \tag{13}$$

for all n, while the function

$$f_{0,\varepsilon}(x) = q_0(x) + \sqrt[3]{\varepsilon}\sum_{n=1}^\infty \theta_n |x_n - x|^2$$

on D reaches its minimum at x_*. Observe that, furthermore, in terms of Theorem 1 of [15] we take $\theta_n = \sqrt[3]{\varepsilon^2}\gamma_n\alpha_n^{-2}$.

To the problem with inequality constraints

$$f_{0,\varepsilon}(x) \to \min,\quad q_j(x) - \delta_j^2 \le 0,\quad j = 1,\dots,m,$$

at x_* apply the second-order necessary conditions (see [17, Theorem 2.1]). By this theorem, there exist Lagrange multipliers $\bar\lambda_\varepsilon = (\lambda_\varepsilon^0,\dots,\lambda_\varepsilon^m) \in \bar\Lambda$ satisfying Lagrange's equation

$$\frac{\partial\mathscr{L}_\varepsilon}{\partial x}(x_*, \bar\lambda_\varepsilon) = 0, \tag{14}$$

the complementary slackness conditions

$$\lambda_\varepsilon^j\big(q_j(x_*)-\delta_j^2\big)=0,\quad j=1,\dots,m, \tag{15}$$

and the second-order conditions

$$\operatorname{ind}\bigg(\frac{\partial^2\mathscr{L}_\varepsilon}{\partial x^2}(x_*,\bar\lambda_\varepsilon)\bigg)\le m. \tag{16}$$

Here

$$\mathscr{L}_\varepsilon(x,\bar\lambda)=\lambda^0 f_{0,\varepsilon}(x)+\sum_{j=1}^m\lambda^j\big(q_j(x)-\delta_j^2\big).$$

By construction, we can express $f_{0,\varepsilon}$ as a converging series of quadratic forms. Moreover, both this series and the series of derivatives converge uniformly on every bounded set, while q_j are quadratic forms. This yields

$$\langle f_{0,\varepsilon}'(x),x\rangle=2f_{0,\varepsilon}(x),\quad \langle q_j'(x),x\rangle=2q_j(x),\quad j=0,\dots,m.$$

Therefore, multiplying (14) by x_* and accounting for the complementary slackness conditions (15), we obtain

$$\lambda_\varepsilon^0\bigg(q_0(x_*)+\sqrt[3]{\varepsilon}\sum_{n=1}^\infty\theta_n|x_n-x_*|^2\bigg)=-\sum_{j=1}^m\lambda_\varepsilon^j\delta_j^2. \tag{17}$$

Passing in the second inequality of (13) to the limit as $n\to\infty$, we find $|x_*-w|\le\sqrt[3]{\varepsilon}$, which by (13) and the triangle inequality yields $|x_n-x_*|\le 2\sqrt[3]{\varepsilon}$ for all n. Passing in the third inequality in (13) to the limit as $n\to\infty$, we obtain $q_0(x_*)\le q_0(w)\le\kappa+\varepsilon$, whence $|q_0(x_*)-\kappa|\le\varepsilon$ because $x_*\in D$ and so $q_0(x_*)\ge\kappa$. From the resulting inequalities with $\lambda_\varepsilon^0\le 1$ we deduce that

$$\bigg|\lambda_\varepsilon^0\bigg(q_0(x_*)+\sqrt[3]{\varepsilon}\sum_{n=1}^\infty\theta_n|x_n-x_*|^2\bigg)-\lambda_\varepsilon^0\kappa\bigg|\le 5\varepsilon. \tag{18}$$

Let us take $\mu=(\mu^1,\dots,\mu^m)$ with arbitrary $\mu_j\ge 0$ and show that for $\bar\mu_\varepsilon=(\lambda_\varepsilon^0,\mu)$ we have

$$\inf_{x\in X}\mathscr{L}_0(x,\bar\mu_\varepsilon)\le\lambda_\varepsilon^0\kappa+5\varepsilon. \tag{19}$$

Indeed, since $x_*\in D$, it follows that $\mu^j\big(q_j(x_*)-\delta_j^2\big)\le 0$ for $j=1,\dots,m$, and consequently $\mathscr{L}_\varepsilon(x_*,\bar\mu_\varepsilon)\le\lambda_\varepsilon^0 f_{0,\varepsilon}(x_*)$. Therefore, the obvious inequality $f_{0,\varepsilon}(x)\ge q_0(x)$, valid for all $x\in X$, yields

$$\inf_{x\in X}\mathscr{L}_0(x,\bar\mu_\varepsilon)\le\inf_{x\in X}\mathscr{L}_\varepsilon(x,\bar\mu_\varepsilon)\le\mathscr{L}_\varepsilon(x_*,\bar\mu_\varepsilon)\le\lambda_\varepsilon^0 f_{0,\varepsilon}(x_*).$$

Hence, (18) implies (19).

Assume henceforth that ε^{-1} takes only positive integer values. Extracting from the bounded sequence $\{\bar\lambda_\varepsilon\}$ of $(m+1)$-dimensional vectors a subsequence, assume that $\bar\lambda_\varepsilon\to\bar\lambda$ for some vector $\bar\lambda=(\lambda^0,\dots,\lambda^m)$. It is obvious that $\bar\lambda\in\bar\Lambda$.

Let us study the family of quadratic forms $\frac{\partial^2\mathscr{L}_\varepsilon}{\partial x^2}(x,\bar\lambda)$. As noted above, we can express the function $f_{0,\varepsilon}$ as a converging series of quadratic forms, while q_j are quadratic forms. Therefore, the quadratic form $\frac{\partial^2\mathscr{L}_\varepsilon}{\partial x^2}(x,\bar\lambda)$ is independent of x and depends only on $\bar\lambda$ and ε. But if a sequence of quadratic forms whose indices are bounded above by some number m converges uniformly on the unit ball to some quadratic form then its index satisfies the same bound. This claim follows straightforwardly from Theorem 2.3 of [17]. Therefore, by (16) the index of the quadratic form $\frac{\partial^2\mathscr{L}_0}{\partial x^2}(x,\bar\lambda)=\lambda^0q_0+\dots+\lambda^mq_m$ is at most m. Consequently, by condition $\mathscr{A}$ this quadratic form is nonnegative definite.

By (17) and (18), we have $\left|\lambda^0_\varepsilon\kappa + \sum_{j=1}^m \lambda^j_\varepsilon\delta_j^2\right| \le 5\varepsilon$. Passing here to the limit as $\varepsilon \to 0$, we infer that

$$\lambda^0\kappa = -\sum_{j=1}^m \lambda^j\delta_j^2. \tag{20}$$

But then $\lambda^0 > 0$ because all λ^j are nonnegative and not simultaneously vanishing, while all δ_j are positive by assumption. Taking into account the positive homogeneity of the resulting relations with respect to $\bar\lambda$ and dividing them by λ^0, without loss of generality we assume that $\lambda^0 = 1$. The quadratic form $q_0 + \lambda^1 q_1 + \dots + \lambda^m q_m$ is nonnegative definite, while the minimum of each nonnegative definite quadratic form equals zero. Therefore,

$$\min_{x\in X} L(x,\lambda) = -\sum_{j=1}^m \lambda^j\delta_j^2,$$

whence by (20) we obtain (9). $\square$

Let us formulate the sufficient conditions that ensure the fulfillment of condition $\mathscr{A}$.

Lemma 2. *Take a dense linear subspace $\widetilde X$ of X. Suppose that for every $h \in \widetilde X$ with $h \ne 0$ there exists a linear operator $B = B_h : \widetilde X \to \widetilde X$ such that for all $j = 0, 1, \dots, m$ we have*

(1) $q_j(B^k h) \le q_j(h)$ *for* $k = 1, \dots, m$;

(2) $\langle Q_j B^{k_1} h, B^{k_2} h\rangle = 0$ *for* $0 \le k_1 < k_2 \le m$.

Then the quadratic forms $q_j(x)$ for $j = 0, 1, \dots, m$ satisfy condition $\mathscr{A}$.

Proof. Suppose that there exist $\lambda^j \ge 0$ for $j = 0, 1, \dots, m$, for which the quadratic form

$$q = \lambda^0 q_0 + \lambda^1 q_1 + \dots + \lambda^m q_m$$

is not nonnegative definite. Then there exists $h \in X$ with $q(h) < 0$. Since $\widetilde X$ is dense in X, we may assume that $h \in \widetilde X$. Let us show that the index of q is greater than m. Consider the system of vectors $x_j = B^j h$ for $j = 0, 1, \dots, m$. Verify that for all $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_m) \ne 0$ the vector

$$x = \sum_{k=0}^m \alpha_k x_k$$

satisfies $q(x) < 0$. Properties 1 and 2 yield

$$\begin{aligned} q_j(x) &= \left\langle Q_j \sum_{k=0}^m \alpha_k x_k, \sum_{k=0}^m \alpha_k x_k \right\rangle = \sum_{k=0}^m \alpha_k^2 \langle Q_j x_k, x_k\rangle \\ &= \sum_{k=0}^m \alpha_k^2 \langle Q_j B^k h, B^k h\rangle \le \sum_{k=0}^m \alpha_k^2 q_j(h) = q_j(h)\sum_{k=0}^m \alpha_k^2. \end{aligned}$$

Hence,

$$q(x) = \sum_{j=0}^m \lambda^j q_j(x) \le \sum_{k=0}^m \alpha_k^2 \left(\sum_{j=0}^m \lambda^j q_j(h)\right) = q(h)\sum_{k=0}^m \alpha_k^2 < 0,$$

and so $x \ne 0$. Therefore, $x_0, x_1, \dots, x_m$ are linearly independent, while q is negative definite on their linear span of dimension $m + 1$. Thus, the index of this form exceeds m. $\square$

Remark. As B_h, it is often convenient to take $B_h = A^{n(h)}$, where $A : \widetilde X \to \widetilde X$ is a given linear operator, while $n(h)$ for each h takes positive integer values.

For instance, take as X the Hilbert space of pairs of vector functions $w(t) = (\xi(t), u(t))$ for $t \in \mathbb{R}$, where $u(\cdot)$ is a Lebesgue measurable m-dimensional function $u(\cdot) : \mathbb{R} \to \mathbb{R}^m$ whose squared absolute value is summable on $\mathbb{R}$ (denote the set of these functions by $L_2^m(\mathbb{R})$), while the absolutely continuous n-dimensional function $\xi(\cdot)$ is a solution of the equation

$$\dot\xi = D\xi + Eu(t); \tag{21}$$

furthermore, $\xi(\cdot) \in L_2^n(\mathbb{R})$. Here D and E are given real matrices of appropriate sizes. Define the quadratic forms q_j as

$$q_j(w(\cdot)) = \int\limits_{\mathbb{R}} (\langle G_j\xi(t), \xi(t)\rangle + 2\langle Q_j\xi(t), u(t)\rangle + \langle R_j u(t), u(t)\rangle)\, dt, \tag{22}$$

where G_j, Q_j, and R_j are given matrices of appropriate sizes.

It is known, see [18], that the subspace $\widetilde{X}$ consisting of compactly-supported vector functions $w(\cdot) \in X$ is dense in X. As the operator A, take the time-shift by 1; i.e., $Aw(t) = w(t-1)$. By Lemma 2 and the remark following it, in this example condition $\mathscr{A}$ holds. The same arguments remain valid if in (22) the integration, instead of the axis $\mathbb{R}$, goes over the positive ray $\mathbb{R}^+$ or the negative ray $\mathbb{R}^-$, while the trajectory $\xi(\cdot)$ obeys the extra condition $\xi(0) = 0$. Furthermore, in the first case in the definition of the space X we have to assume additionally that $u(t) = 0$ and $\xi(t) = 0$ for all $t \le 0$, whereas in the second case that $u(t) = 0$ and $\xi(t) = 0$ for all $t \ge 0$, while taking as A the time-shift operator $Aw(t) = w(t+1)$.

Proceed to a more general construction. Consider the measure space (T, Σ, μ), where T is a nonempty set, Σ is a σ-algebra of subsets of T, called measurable, and μ is a nonnegative countably additive σ-finite set function. The latter means that in T there exists a sequence $T_n \in \Sigma$ of measurable subsets such that $\mu(T_n) < \infty$ for all n and $\bigcup_n T_n = T$.

Take a Hilbert space Y and consider the $L_2(T, \mu, Y)$ space consisting of measurable mappings $f : T \to Y$ the square of whose absolute value $|f|^2 = \langle f, f\rangle$ is summable on T. The inner product on it is defined in the usual way. It is known [19] that $L_2(T, \mu, Y)$ is itself a Hilbert space.

In the family of measurable sets select a subfamily $\widetilde{\Sigma} \subset \Sigma$, assuming that $\mu(e) < \infty$ for all $e \in \widetilde{\Sigma}$, and that this subfamily is closed under finite unions; i.e., $e_1, e_2 \in \widetilde{\Sigma}$ implies $e_1 \cup e_2 \in \widetilde{\Sigma}$.

On $L_2(T, \mu, Y)$ consider the quadratic forms

$$q_j(f) = \int\limits_T \langle C_j f(t), f(t)\rangle\, d\mu, \quad j = 0, \dots, m. \tag{23}$$

Here $C_j : Y \to Y$ are given symmetric linear operators.

Theorem 2. *Given a closed linear space $X \subset L_2(T, \mu, Y)$, take a dense linear subspace $\widetilde{X}$ of X and assume that there exists a measurable mapping $\varphi : T \to T$ (i.e., the set $\varphi^{-1}(e)$ is measurable for every measurable $e \in \Sigma$) satisfying the conditions:*

(1) μ is invariant under φ on $\widetilde{\Sigma}$ in the sense that $\varphi^{-1}(e) \in \widetilde{\Sigma}$ and $\mu(e) = \mu(\varphi^{-1}(e))$ for every $e \in \widetilde{\Sigma}$;

(2) for every measurable $e \in \widetilde{\Sigma}$ there exists a positive integer $n = n(e)$ such that $\mu(e \cap \varphi^{-jn}(e)) = 0$ for each $j = 1, \dots, m$ (here φ^{-s} is the sth iteration of the inverse mapping φ^{-1});

(3) the linear operator of composition $f \to f \circ \varphi = f(\varphi)$ keeps invariant the subspace $\widetilde{X}$ and the support of every $f \in \widetilde{X}$ lies in $\widetilde{\Sigma}$, i.e., $\{t \in T : f(t) \ne 0\} \in \widetilde{\Sigma}$.

Then the Lagrange function minimality condition holds for (5) *in which the quadratic forms q_j are defined in* (23).

Observe that, by the first two assumptions on the measure μ, if it is nonzero then $\mu(T) = \infty$.

Proof. Take $h(\cdot) \in \widetilde{X}$ and construct the corresponding linear operator B_h satisfying the conditions of Lemma 2.

Denote by $e_0 = \{t \in T : h(t) \ne 0\}$ the support of h. Then $e_0 \in \widetilde{\Sigma}$ by condition 3. By condition 2, there exists a positive integer $n = n(e_0)$ such that $\mu(e_0 \cap \tilde{\varphi}^j(e_0)) = 0$ for $j = 1, \dots, m$, where $\tilde{\varphi} = \varphi^{-n}$.

Introduce the sets $e_j = \tilde{\varphi}^j(e_0)$ for $j = 1, \dots, m$. Then condition 1 yields $e_j \in \widetilde{\Sigma}$. Moreover, $\mu(e_j \cap e_0) = 0$ for each $j = 1, \dots, m$, as well as $\mu(e_j) = \kappa$, where $\kappa = \mu(e_0) < \infty$ because μ is also invariant under φ^{nj}.

Let us show that

$$\mu(e_{k_1} \cap e_{k_2}) = 0 \quad \forall\, k_1, k_2 : 0 \le k_1 < k_2 \le m. \tag{24}$$

For $k_1 = 0$ these equalities are noted above. Assume that $k_1 \geq 1$ and verify that then

$$\mu(e_{k_1} \cup e_{k_2}) = \mu(e_{k_1-1} \cup e_{k_2-1}). \tag{25}$$

Indeed, since the image of the union of two sets equals the union of their images,

$$\tilde{\varphi}(e_{k_1} \cup e_{k_2}) = \tilde{\varphi}(e_{k_1}) \cup \tilde{\varphi}(e_{k_2}) = e_{k_1-1} \cup e_{k_2-1}$$
$$\Longrightarrow \mu(e_{k_1} \cup e_{k_2}) = \mu(\tilde{\varphi}(e_{k_1} \cup e_{k_2})) = \mu(e_{k_1-1} \cup e_{k_2-1}),$$

which proves (25).

From (25), decreasing the positive integer k_1 to zero, we infer that $\mu(e_{k_1} \cup e_{k_2}) = \mu(e_0 \cup e_j)$ for $j = k_2 - k_1$. However, $\mu(e_0 \cap e_j) = 0$ and so $\mu(e_{k_1} \cup e_{k_2}) = 2\kappa$. Moreover, $\mu(e_{k_1}) + \mu(e_{k_2}) = 2\kappa$; since $\mu(e_{k_1} \cap e_{k_2}) = \mu(e_{k_1}) + \mu(e_{k_2}) - \mu(e_{k_1} \cup e_{k_2})$ we obtain (24).

By condition 3, define the linear operator of composition $A : \widetilde{X} \to \widetilde{X}$ as $(Af)(t) = f(\varphi(t))$ for $t \in T$ and the operator $B_h : \widetilde{X} \to \widetilde{X}$ as $B_h = A^n$, where $n = n(e_0)$. By construction, $\big(B_h^k h\big)(t) = h(\varphi^{nk}(t)) = 0$ for almost all $t \notin e_k$. By (24), for all $k_1 < k_2$ and almost all $t \in T$ this yields $\big\langle C_j \big(B_h^{k_1} h\big)(t), \big(B_h^{k_2} h\big)(t)\big\rangle = 0$. The argument directly implies the validity of condition 2 of Lemma 2.

Let us verify condition 1 of Lemma 2. To this end, it suffices to show that

$$\int_T \langle C_j h(t), h(t)\rangle \, d\mu = \int_T \langle C_j h(\varphi^{nk}(t)), h(\varphi^{nk}(t))\rangle \, d\mu \tag{26}$$

for all j and k. Indeed, let us verify that every set $\widetilde{T} \subset T$ of finite measure satisfies

$$\int_{\widetilde{T}} \langle C_j h(t), h(t)\rangle \, d\mu = \int_{\widetilde{T}} \langle C_j h(\varphi^{nk}(t)), h(\varphi^{nk}(t))\rangle \, d\mu. \tag{27}$$

Fix some $\varepsilon > 0$. Take a simple function h_ε such that $|\langle C_j h(t), h(t)\rangle - \langle C_j h_\varepsilon(t), h_\varepsilon(t)\rangle| \leq \varepsilon$ for almost all $t \in T$. Assume that h_ε takes countably many values y_s for $s = 1, 2, \dots$. Then

$$\int_{\widetilde{T}} \langle C_j h_\varepsilon(t), h_\varepsilon(t)\rangle \, d\mu = \sum_{s=1}^{\infty} \langle C_j y_s, y_s\rangle \mu(T_s), \quad T_s = \{t \in \widetilde{T} : h_\varepsilon(t) = y_s\}.$$

Consider the simple function $h_\varepsilon(\varphi^{nk})$. For each s we have

$$\mu(\{t \in \widetilde{T} : h_\varepsilon(\varphi^{nk}(t)) = y_s\}) = \mu(\varphi^{(-nk)}(T_s)) = \mu(T_s).$$

This implies that

$$\int_{\widetilde{T}} \langle C_j h_\varepsilon(t), h_\varepsilon(t)\rangle \, d\mu = \int_{\widetilde{T}} \langle C_j h_\varepsilon(\varphi^{nk}(t)), h_\varepsilon(\varphi^{nk}(t))\rangle \, d\mu,$$

which forces (27) because $\varepsilon > 0$ is arbitrary. The validity of (26) follows because (27) holds for every set $\widetilde{T}$ of finite measure. Thus, condition 1 of Lemma 2 also holds (and moreover, as an equality).

By Lemma 2, the quadratic forms $q_j(x)$ for $x \in X$ and $j = 0, 1, \dots, m$ defined in (23) satisfy condition $\mathscr{A}$. Thus, the validity of Theorem 2 follows from Theorem 1. $\square$

Let us provide several natural examples in which the main assumptions of Theorem 2 are fulfilled. To start with, take $T = \mathbb{R}^d$, the d-dimensional space with Lebesgue measure μ. Then we can take as $\widetilde{\Sigma}$ the family of bounded Lebesgue measurable sets. Take as the mapping φ the translation by an arbitrary fixed nonzero vector $\bar{t} \in \mathbb{R}^d$, i.e., $\varphi(t) = t + \bar{t}$ for $t \in \mathbb{R}^d$.

Now take $T = \mathbb{Z}^d$, where $\mathbb{Z}^d$ is the set of d-dimensional integer vectors z, while μ is a discrete measure such that $\mu(z) = 1$ for all $z \in \mathbb{Z}^d$. Take as $\widetilde{\Sigma}$ the family of bounded subsets of $\mathbb{Z}^d$, while as φ the translation by a nonzero integer vector $\bar{z}$, i.e., $\varphi(z) = z + \bar{z}$ for $z \in \mathbb{Z}^d$.

Let us give one more example. Take the cylinder $T = T_1 \times T_2$, where T_1 is either $\mathbb{R}^d$ or $\mathbb{Z}^d$, while the measure μ_1 defined on T_1 is accordingly either the Lebesgue measure or the discrete measure described above. Take as T_2 a measurable subset of $\mathbb{R}^d$ or $\mathbb{Z}^d$ with the corresponding measure μ_2. Define the measure μ on T as the product measure, i.e., $\mu = \mu_1 \otimes \mu_2$. In this case, as above, take as $\widetilde{\Sigma}$ the family of measurable bounded subsets (this is always the natural choice for $\widetilde{\Sigma}$ whenever T is equipped with a metric). As φ we have to take the translation along the space T_1, i.e., $\varphi(t) = t + (\bar{t}_1, 0)$ for $t = (t_1, t_2) \in T$, where $\bar{t}_1$ is an arbitrary nonzero element of T_1. Observe that in all cases the subspace $\widetilde{X}$ consists of compactly supported functions.

2. Recovery Problems on Assuming the Lagrange Function Minimality Condition

It turns out that when the Lagrange function minimality condition is fulfilled for (5), together with an additional solvability condition, we can explicitly express the error $E(I, \delta)$ of optimal recovery in terms of the corresponding Lagrange multipliers, as well as find an optimal recovery method.

In the direct product of $\widehat{Y} = Y_{k+1} \times \dots \times Y_m$ define the norm in the usual fashion as

$$\|y\|_{\widehat{Y}} = \left(\sum_{j=k+1}^{m} \|y_j\|_{Y_j}^2 \right)^{1/2}.$$

Theorem 3. *Assume that in (5), where q_j for $j = 0, 1, \dots, m$ are defined by (6), the Lagrange function minimality condition holds with Lagrange multipliers λ^j for $j = 1, \dots, m$. Suppose that there exist dense linear subspaces $\widetilde{Y}_j \subseteq Y_j$ for $j = k+1, \dots, m$, and a continuous linear operator $A : Y_{k+1} \times \dots \times Y_m \to Y_0$ such that for all $y = (y_{k+1}, \dots, y_m) \in \widetilde{Y}_{k+1} \times \dots \times \widetilde{Y}_m$ there exists a solution $x_y \in X$ to the equation*

$$\sum_{j=1}^{m} \lambda^j I_j^* I_j x = \sum_{j=k+1}^{m} I_j^* y_j \tag{28}$$

and $Ay = I_0 x_y$. Then the error of optimal recovery equals

$$E(I, \delta) = \left(\sum_{j=1}^{m} \lambda^j \delta_j^2 \right)^{1/2},$$

while the method $\varphi(y) = A\Lambda y$, where $\Lambda y = (\lambda^{k+1} y_{k+1}, \dots, \lambda^m y_m)$, is optimal.

Proof. Lemma 1 and Proposition 1 yield the lower bound

$$E(I, \delta) \geq \left(\sum_{j=1}^{m} \lambda^j \delta_j^2 \right)^{1/2}. \tag{29}$$

To obtain an upper bound, consider the linear space $E = Y_1 \times \dots \times Y_m$ with the inner semiproduct

$$(y^1, y^2)_E = \sum_{j=1}^{m} \lambda^j \big(y_j^1, y_j^2\big)_{Y_j},$$

where $y^1 = \big(y_1^1, \dots, y_m^1\big)$ and $y^2 = \big(y_1^2, \dots, y_m^2\big)$. Put $\widetilde{I}x = (I_1 x, \dots, I_m x)$ and $\tilde{y}_0 = (0, \dots, 0, y_{k+1}, \dots, y_m)$. If $x_{\Lambda y} \in X$ is a solution to (28) then $(\widetilde{I}x_{\Lambda y}, \widetilde{I}x)_E = (\tilde{y}_0, \widetilde{I}x)_E$ for all $x \in X$. This yields

$$\|\widetilde{I}x - \tilde{y}_0\|_E^2 = \|\widetilde{I}x - \widetilde{I}x_{\Lambda y} + \widetilde{I}x_{\Lambda y} - \tilde{y}_0\|_E^2 = \|\widetilde{I}x - \widetilde{I}x_{\Lambda y}\|_E^2 + \|\widetilde{I}x_{\Lambda y} - \tilde{y}_0\|_E^2.$$

Thus,

$$\|\widetilde{I}x - \widetilde{I}x_{\Lambda y}\|_E^2 \leq \|\widetilde{I}x - \tilde{y}_0\|_E^2 = \sum_{j=1}^{k} \lambda^j \|I_j x\|_{Y_j}^2 + \sum_{j=k+1}^{m} \lambda^j \|I_j x - y_j\|_{Y_j}^2 \tag{30}$$

for all $x \in X$.

Suppose that $x \in W$ and $y = (y_{k+1}, \dots, y_m) \in Y_{k+1} \times \dots \times Y_m$ satisfy $\|I_j x - y_j\| \le \delta_j$ for $j = k+1, \dots, m$. Then for every $\varepsilon > 0$ there exists $\tilde{y} = (\tilde{y}_{k+1}, \dots, \tilde{y}_m) \in \widetilde{Y}_{k+1} \times \dots \times \widetilde{Y}_m$ such that $\|y_j - \tilde{y}_j\|_{Y_j} \le \varepsilon$. Therefore,

$$\|I_j x - \tilde{y}_j\|_{Y_j} \le \|I_j x - y_j\|_{Y_j} + \|y_j - \tilde{y}_j\|_{Y_j} \le \delta_j + \varepsilon, \quad j = k+1, \dots, m.$$

Put $z = x - x_{\Lambda\tilde{y}}$. Then (30) implies that

$$\sum_{j=1}^{m} \lambda^j \|I_j z\|_{Y_j}^2 \le \sum_{j=1}^{m} \lambda^j \tilde{\delta}_j^2, \tag{31}$$

where

$$\tilde{\delta}_j = \begin{cases} \delta_j & \text{for } 1 \le j \le k, \\ \delta_j + \varepsilon & \text{for } k+1 \le j \le m. \end{cases}$$

We estimate the error of the method $\varphi(y) = A\Lambda y$ as

$$\|I_0 x - A\Lambda y\|_{Y_0} \le \|I_0 x - A\Lambda\tilde{y}\|_{Y_0} + \|A\Lambda(\tilde{y} - y)\|_{Y_0} \le \|I_0 x - I_0 x_{\Lambda\tilde{y}}\|_{Y_0} + \|A\Lambda\|(m-k)\varepsilon.$$

It is not difficult to verify that for all $a, b > 0$ we have

$$\sup_{\substack{z \in X \\ \sum_{j=1}^{m} \lambda^j \|I_j z\|_{Y_j}^2 \le a^2}} \|I_0 z\|_{Y_0}^2 = \frac{a^2}{b^2} \sup_{\substack{x \in X \\ \sum_{j=1}^{m} \lambda^j \|I_j x\|_{Y_j}^2 \le b^2}} \|I_0 x\|_{Y_0}^2.$$

Using (31) and Proposition 1, we obtain

$$\|I_0 x - I_0 x_{\Lambda\tilde{y}}\|_{Y_0}^2 = \|I_0 z\|_{Y_0}^2 \le \sup_{\substack{z \in X \\ \sum_{j=1}^{m} \lambda^j \|I_j z\|_{Y_j}^2 \le \sum_{j=1}^{m} \lambda^j \tilde{\delta}_j^2}} \|I_0 z\|_{Y_0}^2$$

$$= \frac{\sum_{j=1}^{m} \lambda^j \tilde{\delta}_j^2}{\sum_{j=1}^{m} \lambda^j \delta_j^2} \sup_{\substack{z \in X \\ \sum_{j=1}^{m} \lambda^j \|I_j z\|_{Y_j}^2 \le \sum_{j=1}^{m} \lambda^j \delta_j^2}} \|I_0 z\|_{Y_0}^2 = \sum_{j=1}^{m} \lambda^j \tilde{\delta}_j^2.$$

Since $\varepsilon > 0$ is arbitrary,

$$\|I_0 x - A\Lambda y\|_{Y_0} \le \left(\sum_{j=1}^{m} \lambda^j \delta_j^2 \right)^{1/2}. \tag{32}$$

It follows from (29) and (32) that

$$E(I, \delta) = \left(\sum_{j=1}^{m} \lambda^j \delta_j^2 \right)^{1/2},$$

and the method $\varphi(y) = A\Lambda y$ is optimal. $\square$

Let us make several remarks. Equation (28) arises from the necessity of solving the extremal problem

$$\sum_{j=1}^{k} \lambda^j \|I_j x\|_{Y_j}^2 + \sum_{j=k+1}^{m} \lambda^j \|I_j x - y_j\|_{Y_j}^2 \to \min, \quad x \in X.$$

It appears in the approach to constructing an optimal recovery method that was proposed in [9]. However, in some cases this equation has a solution only for some $y_j \in Y_j$, where $j = k+1, \dots, m$. This situation comes up, for instance, in recovering solutions to the heat equation at a fixed time from noisy solutions to this equation at times [12]. In connection with this, we have to consider solving (28) on the direct product of dense subsets of Y_j for $j = k+1, \dots, m$.

Return now to the general construction of Theorem 2. Assume that T, μ, Y, and X satisfy the conditions of the theorem. Consider problem (2), where X is the same as in Theorem 2, while $Y_0, Y_1, \dots, Y_m$ are Hilbert spaces. Put $C_0 = -I_0^* I_0$ and $C_j = I_j^* I_j$ for $j = 1, \dots, m$. Theorem 2 ensures the fulfillment of the Lagrange function minimality condition with certain Lagrange multipliers $\lambda^j \geq 0$ for $j = 1, \dots, m$, which in particular implies that the quadratic form generated by $C = \lambda^1 C_1 + \dots + \lambda^m C_m$ is nonnegative definite.

Assume in addition that C is strictly positive, i.e., there exists $\varepsilon > 0$ such that $\langle Cx, x \rangle \geq \varepsilon |x|^2$ for all $x \in Y$. Then C is continuously invertible. Therefore, Theorem 3 yields

$$E(I, \delta) = \left(\sum_{j=1}^{m} \lambda^j \delta_j^2 \right)^{1/2},$$

and the method

$$\varphi(y) = I_0(\lambda^1 C_1 + \dots + \lambda^m C_m)^{-1}(\lambda^{k+1} I_{k+1}^* y_{k+1} + \dots + \lambda^m I_m^* y_m)$$

is optimal.

References

1. Smolyak S. A., *On Optimal Recovery of Functions and Functionals over Them* [Russian], Diss. Kand. Fiz.-Mat. Nauk, Moscow Univ., Moscow (1965).
2. Marchuk A. G. and Osipenko K. Yu., "Best approximation of functions specified with an error at a finite number of points," Math. Notes, vol. 17, no. 3, 207–212 (1975).
3. Osipenko K. Yu., "Best approximation of analytic functions from information about their values at a finite number of points," Math. Notes, vol. 19, no. 1, 17–23 (1976).
4. Micchelli C. A. and Rivlin T. J., "A survey of optimal recovery," in: *Optimal Estimation in Approximation Theory* (C. A. Micchelli and T. J. Rivlin; eds.), Plenum Press, New York, 1977, 1–54.
5. Arestov V. V., "Optimal recovery of operators, and related problems," Trudy Mat. Inst. Steklov., vol. 189, 3–20 (1989).
6. Traub J. F. and Woźniakowski H., *A General Theory of Optimal Algorithms*, Academic Press, New York (1980).
7. Plaskota L., *Noisy Information and Computational Complexity*, Cambridge Univ. Press, Cambridge (1996).
8. Osipenko K. Yu., *Optimal Recovery of Analytic Functions*, Nova Sci. Publ., Inc., Huntington and New York (2000).
9. Melkman A. A. and Micchelli C. A., "Optimal estimation of linear operators in Hilbert spaces from inaccurate data," SIAM J. Numer. Anal., vol. 16, 87–105 (1979).
10. Magaril-Il′yaev G. G. and Osipenko K. Yu., "Optimal recovery of functions and their derivatives from Fourier coefficients prescribed with an error," Sb. Math., vol. 193, no. 3, 387–407 (2002).
11. Osipenko K. Yu., "The Hardy–Littlewood–Pólya inequality for analytic functions in Hardy–Sobolev spaces," Sb. Math., vol. 197, no. 3, 315–334 (2006).
12. Magaril-Il′yaev G. G. and Osipenko K. Yu., "Optimal recovery of the solution of the heat equation from inaccurate data," Sb. Math., vol. 200, no. 5, 665–682 (2009).
13. Arutyunov A. V., "Milyutin's theorem in linear-quadratic optimal control problems," Differ. Equ., vol. 37, no. 11, 1627–1630 (2001).
14. Arutyunov A. V., "Lagrange principle in quadratic optimal control problems with infinite horizon," Differ. Equ., vol. 45, no. 11, 1595–1601 (2009).
15. Ioffe A. D. and Tikhomirov V. M., "Some remarks on variational principles," Math. Notes, vol. 61, no. 2, 248–253 (1997).
16. Bobylev N. A., Emel′yanov S. V., and Korovin S. K., *Geometric Methods in Variational Problems* [Russian], Magistr, Moscow (1998).
17. Arutyunov A. V., "Smooth abnormal problems in extremum theory and analysis," Russian Math. Surveys, vol. 67, no. 3 (405), 403–457 (2012).
18. Arutyunov A. V., "Approximation to solutions of linear control systems by compactly supported solutions," Differ. Equ., vol. 51, no. 6, 792–797 (2015).
19. Yosida K., *Functional Analysis*, Springer-Verlag, Berlin (1994).

A. V. Arutyunov
Peoples' Friendship University of Russia, Moscow, Russia
E-mail address: arutyunov@cs.msu.ru

K. Yu. Osipenko
Moscow State University
Institute for Information Transmission Problems, Moscow, Russia
E-mail address: kosipenko@yahoo.com