Adaptive and Nonadaptive Recovery

K. Yu. Osipenko[1]

Moscow State University,
Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow

Abstract

The article is devoted to the comparison of adaptive and nonadaptive recovery methods. We obtain some conditions when adaptive does not help for recovery of linear operators. For recovery of linear functionals we give necessary and sufficient conditions for the existence of an optimal nonadaptive linear method which gives the same error as optimal adaptive methods.

Keywords: optimal recovery, adaptive and nonadaptive methods, adaptive radius of information
2010 MSC: 41A65, 41A45, 41A46, 49N30

1. Introduction

Let X be a linear space, Z be a normed linear space, $T\colon X \to Z$ be a linear operator and $W \subset X$. We consider the problem of recovery of T on the set $W \subset X$ by noisy information about elements from W. Suppose that there is a family $\mathcal{I}$ of linear information operators $I_p\colon X \to Y$, $p \in \Omega$, where Ω is some set and Y is a normed linear space. All linear spaces are considered over the field of real or complex numbers. Let us choose n parameters $p_1, \ldots, p_n \in \Omega$ and assume that for all $x \in W$ we know $y_1, \ldots, y_n \in Y$ such that

$$\|I_{p_j}x - y_j\|_Y \le \delta, \quad \delta \ge 0.$$

As recovery methods we consider all possible mappings $\varphi\colon Y^n \to Z$. The error of a method φ is defined as follows

$$e_n(T, W, \mathcal{I}, \delta, p, \varphi) = \sup_{\substack{x \in W,\ y=(y_1,\ldots,y_n) \in Y^n \\ \|I_{p_j}x - y_j\|_Y \le \delta,\ j=1,\ldots,n}} \|Tx - \varphi(y)\|_Z;$$

here $p = (p_1, \ldots, p_n)$. The quantity

$$E_n(T, W, \mathcal{I}, \delta) = \inf_{p_1,\ldots,p_n \in \Omega} \inf_{\varphi\colon Y^n \to Z} e_n(T, W, \mathcal{I}, \delta, p, \varphi)$$

Email address: kosipenko@yahoo.com (K. Yu. Osipenko)
[1]The research was carried out with the financial support of the Russian Foundation for Basic Research (grant no. 17-01-00649)

is called the optimal recovery error. Parameters $\widehat{p}_1, \ldots, \widehat{p}_n$ and the method $\widehat{\varphi}$ we call optimal if

$$e_n(T, W, \mathcal{I}, \delta, \widehat{p}, \widehat{\varphi}) = E_n(T, W, \mathcal{I}, \delta), \quad \widehat{p} = (\widehat{p}_1, \ldots, \widehat{p}_n).$$

One of the typical examples of this setting is the problem of optimal integration. Let W be a class of functions defined on the segment $[a, b]$. Set

$$Tx = \int_a^b x(t)\,dt, \quad I_{t_j} x = x(t_j), \quad t_j \in [a, b],\ j = 1, \ldots, n.$$

Then the problem stated above becomes the problem of choosing optimal method of integration and optimal nodes of integration in which the function should be measured. Quite a lot of research is devoted to this problem (see, for example, [6]).

In practice, quite often a different approach is used, in which subsequent nodes of integration are defined on the base of information about function values in the previous nodes. Such algorithms are called adaptive (or sequential). In this regard, it makes sense to consider a slightly different (more general) setting of the optimal recovery problem.

Let the first parameter p_1 and functions

$$\varphi_j \colon \Omega^{j-1} \times Y^{j-1} \to Z, \quad j = 2, \ldots, n,$$

are given. The following parameters are determined by formulas

$$p_j = \omega_j(p_1, \ldots, p_{j-1}, y_1, \ldots, y_{j-1}), \quad j = 2, \ldots, n, \tag{1}$$

where y_j are approximate values of $I_{p_j} x$, that is, elements of Y such that

$$\|I_{p_j} x - y_j\|_Y \le \delta, \quad j = 1, \ldots, n.$$

As recovery methods we consider all possible mappings

$$\varphi \colon \Omega^n \times Y^n \to Z.$$

Fix $x \in W$. Set

$$\Lambda(x) = \{\, (y_1, \ldots, y_n) : \|I_{p_j} x - y_j\|_Y \le \delta,\ j = 1, \ldots, n \,\},$$

where p_1 is given and p_j, $j = 2, \ldots, n$, are defined by (1).

The error of a method φ is defined as follows

$$e_n^a(T, W, \mathcal{I}, \delta, p_1, \omega, \varphi) = \sup_{\substack{x \in W \\ (y_1, \ldots, y_n) \in \Lambda(x)}} \|Tx - \varphi(p, y)\|_Z;$$

here $\omega = (\omega_2, \ldots, \omega_n)$. The value

$$E_n^a(T, W, \mathcal{I}, \delta) = \inf_{p_1, \omega, \varphi} e_n(T, W, \mathcal{I}, \delta, p_1, \omega, \varphi) \tag{2}$$

is called the error of optimal adaptive recovery.

Adaptive methods have a much richer structure and it is natural to expect that the error of adaptive recovery will be less than for nonadaptive recovery. This is what happens in a number of problems, such as calculating the root of a function or finding the maximum for a unimodal function (note that in these examples, the operator to be recovered is not linear). However, it turns out that there are cases when adaptive recovery does not give a win. In such cases, instead of complicated adaptive methods, it makes sense to use more simpler nonadaptive methods. In [1] it was proved that adaptive methods do not help when T is a linear functional and W is convex and centrally-symmetric set. We prove a more general result, namely, we show that under some condition the corresponding statement remains valid for linear operators. Note that in general case adaptive methods can give a smaller error for recovery of linear operators on convex and centrally-symmetric sets. The example of such situation was constructed in [3].

In the present paper we also give necessary and sufficient conditions for the existence of an optimal nonadaptive linear method which gives the same error as optimal adaptive methods for recovery of linear functionals.

In problem (2) we use mappings

$$\Phi(y) = (p_1, \omega_2(p_1, y_1), \ldots, \omega_n(p_1, \ldots, p_{n-1}, y_1, \ldots, y_{n-1}))$$

which for every $y = (y_1, \ldots, y_n)$ give the values of parameters $p = (p_1, \ldots, p_n)$ by formula $p = \Phi(y)$. Let us extend problem (2) by using arbitrary mappings $\Phi(y)$.

We will consider the following problem of optimal adaptive recovery. Let X, Z, T, and W be as above and $\mathcal{Y}$ be a linear space. Assume that there is a set of parameters $\mathcal{P}$ and for every $p \in \mathcal{P}$ the multivalued mapping $F_p \colon W \to \mathcal{Y}$ is assigned (that is, for every $x \in W$ $F_p(x)$ is a set from $\mathcal{Y}$). The set

$$\operatorname{gr} F_p := \{\, (x, y) : x \in W,\ y \in F_p(x) \,\}$$

is called the graph of F_p. Denote by $\mathcal{F}$ the set of all mappings F_p.

Suppose also that there is a mapping $\Phi \colon \mathcal{Y} \to \mathcal{P}$. Define the mapping G_Φ by its graph

$$\operatorname{gr} G_\Phi = \{\, (x, y) \in W \times \mathcal{Y} : (x, y) \in \operatorname{gr} F_{\Phi(y)} \,\}.$$

As recovery methods of T we consider all possible mappings $\varphi \colon \mathcal{Y} \to Z$ (we do not mark here the dependence of parameters, since the parameter value itself is determined by y).

The error of a method φ is defined as follows

$$e^a(T, \mathcal{F}, \Phi, \varphi) = \sup_{(x,y) \in \operatorname{gr} G_\Phi} \|Tx - \varphi(y)\|_Z,$$

and the value

$$E^a(T, \mathcal{F}) = \inf_{\Phi \colon \mathcal{Y} \to \mathcal{P}} \inf_{\varphi \colon \mathcal{Y} \to Z} e^a(T, \mathcal{F}, \Phi, \varphi)$$

is called the error of optimal adaptive recovery.

By the nonadaptive problem of recovery for the considered case we mean the problem of finding the value

$$E(T,\mathcal{F})=\inf_{p\in\mathcal{P}} e(T,\mathcal{F},p), \tag{3}$$

where

$$\begin{aligned} e(T,\mathcal{F},p) &= \inf_{\varphi\colon \mathcal{Y}\to Z} e(T,\mathcal{F},p,\varphi),\\ e(T,\mathcal{F},p,\varphi) &= \sup_{(x,y)\in \operatorname{gr} F_p} \|Tx-\varphi(y)\|_Z, \end{aligned}$$

and also the problem of finding parameters and methods for which the lower bound is attained.

If $\Phi(y)=p$ for all $y\in\mathcal{Y}$, then

$$e^a(T,\mathcal{F},\Phi,\varphi)=e(T,\mathcal{F},p,\varphi).$$

Therefore,

$$E^a(T,\mathcal{F})\le E(T,\mathcal{F}). \tag{4}$$

2. Some relations between adaptive and nonadaptive recovery errors

Set

$$F_p^{-1}(y)=\{\,x\in W : (x,y)\in \operatorname{gr} F_p\,\}.$$

Put

$$D(p)=\sup_{x\in F_p^{-1}(0)}\|Tx\|_Z.$$

Theorem 1. *Assume that for all $p\in\mathcal{P}$ the sets $\operatorname{gr} F_p$ are centrally-symmetric and contain 0. Then*

$$E^a(T,\mathcal{F})\ge \inf_{p\in\mathcal{P}} D(p).$$

Proof. It follows from the definition of $E^a(T,\mathcal{F})$ that for any $\varepsilon>0$ there exist $\widehat{\Phi}$ and $\widehat{\varphi}$ such that

$$e^a(T,\mathcal{F},\widehat{\Phi},\widehat{\varphi})\le E^a(T,\mathcal{F})+\varepsilon.$$

Consider the value $D(\widehat{\Phi}(0))$. For any $\eta>0$ there exists a $\widehat{x}\in F^{-1}_{\widehat{\Phi}(0)}(0)$ such that

$$\|T\widehat{x}\|_Z\ge D(\widehat{\Phi}(0))-\eta.$$

Since the set $\operatorname{gr} F_{\widehat{\Phi}(0)}$ is centrally-symmetric $-\widehat{x}\in F^{-1}_{\widehat{\Phi}(0)}(0)$. We have

$$\begin{aligned} 2(D(\widehat{\Phi}(0))-\eta) &\le 2\|T\widehat{x}\|_Z=\|T\widehat{x}-\widehat{\varphi}(0)-(T(-\widehat{x})-\widehat{\varphi}(0))\|_Z\\ &\le \|T\widehat{x}-\widehat{\varphi}(0)\|_Z+\|T(-\widehat{x})-\widehat{\varphi}(0)\|_Z\le 2e^a(T,\mathcal{F},\widehat{\Phi},\widehat{\varphi})\\ &\le 2(E^a(T,\mathcal{F})+\varepsilon). \end{aligned}$$

This implies that

$$D(\widehat{\Phi}(0) \le E^a(T, \mathcal{F}).$$

Consequently,

$$\inf_{p\in\mathcal{P}} D(p) \le E^a(T, \mathcal{F}).$$

□

Lemma 1. *Assume that for all $p \in \mathcal{P}$ the sets $\operatorname{gr} F_p$ are convex and centrally-symmetric. Then for all $p \in \mathcal{P}$*

$$e(T, \mathcal{F}, p) \le 2D(p).$$

Proof. Denote by $\operatorname{Pr}_{\mathcal{Y}} \operatorname{gr} F_p$ the set of $y \in \mathcal{Y}$ for which $F_p^{-1}(y) \ne \emptyset$. If $x_1, x_2 \in F_p^{-1}(y)$ then in view of convexity and centrally-symmetry of $\operatorname{gr} F_p$ $\;h = (x_1 - x_2)/2 \in F_p^{-1}(0)$. Set

$$\varphi_0(y) = \begin{cases} Tx(y), & y \in \operatorname{Pr}_{\mathcal{Y}} \operatorname{gr} F_p, \\ 0, & y \notin \operatorname{Pr}_{\mathcal{Y}} \operatorname{gr} F_p, \end{cases}$$

where $x(y)$ is any element from $F_p^{-1}(y)$. Then

$$e(T, \mathcal{F}, p, \varphi_0) = \sup_{y\in \operatorname{Pr}_{\mathcal{Y}} \operatorname{gr} F_p} \; \sup_{x\in F_p^{-1}(y)} \|Tx - \varphi_0(y)\|_Z \le 2\|Th\|_Z,$$

where $h = (x - x(y))/2 \in F_p^{-1}(y)$. Consequently, for all $p \in \mathcal{P}$

$$e(T, \mathcal{F}, p) \le e(T, \mathcal{F}, p, \varphi_0) \le 2D(p).$$

□

In view of (4) from Theorem 1 and Lemma 1 we have

Theorem 2. *If the sets $\operatorname{gr} F_p$ are convex and centrally-symmetric for all $p \in \mathcal{P}$, then*

$$\frac{1}{2} E(T, \mathcal{F}) \le E^a(T, \mathcal{F}) \le E(T, \mathcal{F}).$$

A result similar to Theorem 2 for a less general setting was proved in [2] and [8] (see also [7]).

Theorem 3. *Assume that for all $p \in \mathcal{P}$ the sets $\operatorname{gr} F_p$ are centrally-symmetric, contain 0, and the equality*

$$e(T, \mathcal{F}, p) = D(p) \tag{5}$$

holds. Then

$$E^a(T, \mathcal{F}) = E(T, \mathcal{F}). \tag{6}$$

Proof. It follows from Theorem 1 that

$$E(T,\mathcal{F}) = \inf_{p\in\mathcal{P}} D(p) \le E^a(T,\mathcal{F}).$$

Taking into account (4) we obtain (6). $\square$

Thus, in the case when condition (5) holds optimal nonadaptive methods give the same error as optimal adaptive methods. In other words, in this case adaption does not help.

Condition (5) is valid when T is a linear functional (see [4]). But in some cases this condition may be also valid for linear operators. We give one of such examples.

Let $\mathbb{T}$ be the interval $[-\pi,\pi]$ with identified endpoints. Denote by $\mathcal{W}_2^r(\mathbb{T})$ the Sobolev space of 2π-periodic functions with absolutely continuous the $(r-1)$st derivative such that the rth derivative from $L_2(\mathbb{T})$. Set

$$W_2^r(\mathbb{T}) = \{\, x(\cdot) \in \mathcal{W}_2^r(\mathbb{T}) : \|x^{(r)}(\cdot)\|_{L_2(\mathbb{T})} \le 1 \,\}.$$

Consider the problem of optimal recovery of the mth derivative, $0 \le m < r$, for functions from the class $W_2^r(\mathbb{T})$ by their inaccurate Fourier coefficients.

Let $A \subset \mathbb{Z}_+$ and $B \subset \mathbb{N}$ be finite sets (one of them may be empty). Assume that for every function $x(\cdot) \in W_2^r(\mathbb{T})$ instead of exact values of the Fourier coefficients a_k, $k \in A$ and b_k, $k \in B$, we know their approximate values $\{\widehat{a}_k\}_{k\in A}$ and $\{\widehat{b}_k\}_{k\in B}$ such that

$$|a_k - \widehat{a}_k| \le \delta, \quad k \in A, \quad |b_k - \widehat{b}_k| \le \delta, \quad k \in B.$$

Denote by l_∞^s the space $\mathbb{R}^s$ with the norm $\|y\|_\infty = \max_{0\le j\le s-1} |y_j|$, where $y = (y_0, y_1, \dots, y_{s-1})$. Let

$$A = \{\, k_0, k_1, \dots, k_{N_1} \,\}, \quad B = \{\, l_1, \dots, l_{N_2} \,\},$$

where $k_0 < \ldots < k_{N_1}$, $l_1 < \ldots < l_{N_2}$ and $\operatorname{card} A + \operatorname{card} B = N_1 + 1 + N_2 = s$. Put

$$F_{A,B}x(\cdot) = (a_{k_0}, a_{k_1}, \dots, a_{k_{N_1}}, b_{l_1}, \dots, b_{l_{N_2}}).$$

Then we may say that we know the vector $y \in l_\infty^s$ such that

$$\|F_{A,B}x(\cdot) - y\|_\infty \le \delta.$$

In this case the error of a method $\varphi\colon \mathbb{R}^s \to L_2(\mathbb{T})$ is defined as follows

$$e(D^m, W_2^r(\mathbb{T}), F_{A,B}, \delta, \varphi) = \sup_{\substack{x(\cdot)\in W_2^r(\mathbb{T}),\ y\in l_\infty^s \\ \|F_{A,B}x(\cdot)-y\|_\infty \le \delta}} \|x^{(m)}(\cdot) - \varphi(y)(\cdot)\|_{L_2(\mathbb{T})},$$

where $D^m x(\cdot) = x^{(m)}(\cdot)$. The quantity

$$e(D^m, W_2^r(\mathbb{T}), F_{A,B}, \delta) = \inf_{\varphi\colon l_\infty^s \to L_2(\mathbb{T})} e(D^m, W_2^r(\mathbb{T}), F_{A,B}, \delta, \varphi)$$

is called the error of optimal recovery, and any method $\widehat{\varphi}$ for which this infimum is attained is called an optimal method of recovery.

We are interesting in the problem of choosing the parameter $p = (A, B)$, $\operatorname{card} A + \operatorname{card} B \le N$, so that the error of optimal recovery is minimal. In this problem we would like to consider both non-adaptive methods and adaptive ones. In terms of the general setting here $X = \mathcal{W}_2^r(\mathbb{T})$, $W = W_2^r(\mathbb{T})$, $\mathcal{Y} = \mathbb{R}^N$, $T = D^m$, $Z = L_2(\mathbb{T})$,

$$\begin{gathered}\mathcal{P} = \mathcal{P}_N = \{\, (A, B) \in \mathbb{Z}_+ \times \mathbb{N} : \operatorname{card} A + \operatorname{card} B \le N \,\},\\ F_p x(\cdot) = \{\, y \in l_\infty^s : \|F_{A,B}x(\cdot) - y\|_\infty \le \delta \,\}, \quad s = \operatorname{card} A + \operatorname{card} B.\end{gathered}$$

In [5] it was proved that for all $p = (A, B) \in \mathcal{P}_N$

$$e(D^m, \mathcal{F}, p) = e(D^m, W_2^r(\mathbb{T}), F_{A,B}, \delta) = \sup_{\substack{x(\cdot) \in W_2^r(\mathbb{T}) \\ \|F_{A,B}x(\cdot)\|_\infty \le \delta}} \|x^{(m)}(\cdot)\|_{L_2(\mathbb{T})}.$$

Thus, any adaptive methods (for example, those that choose the number of each next Fourier coefficient based on the values of the previously calculated coefficients) do not provide an advantage over the optimal non-adaptive method.

By virtue of the above, Theorem 4 of [5] implies the following result.

Theorem 4. *Set*

$$\chi_m = \begin{cases} 1, & m = 0, \\ 0, & m \in \mathbb{N}, \end{cases} \quad \widehat{s} = \max\left\{ s \in \mathbb{Z}_+ : 2\delta^2 \sum_{k=0}^{s} k^{2r} < 1, \ s \le [(N - \chi_m)/2] \right\}$$

($[a]$ is the integer part of a). Then

$$\begin{aligned} E^a(D^m, \mathcal{F}) &= E(D^m, \mathcal{F}) \\ &= \sqrt{\frac{1}{(\widehat{s}+1)^{2(r-m)}} + \frac{\delta^2}{2}\chi_m + 2\delta^2 \sum_{k=1}^{\widehat{s}} k^{2m} \left(1 - \left(\frac{k}{\widehat{s}+1}\right)^{2(r-m)}\right)}, \end{aligned}$$

the parameter $\widehat{p} = (\widehat{A}, \widehat{B})$, where for $m = 0$

$$\widehat{A} = (0, 1, \dots, \widehat{s}), \quad \widehat{B} = (1, \dots, \widehat{s}),$$

and for $m > 0$

$$\widehat{A} = (1, \dots, \widehat{s}), \quad \widehat{B} = (1, \dots, \widehat{s}).$$

The method

$$\begin{aligned} \widehat{\varphi}(y)(t) &= \frac{y_0}{2}\chi_m \\ &+ \sum_{k=1}^{\widehat{s}} \left(1 - \left(\frac{k}{\widehat{s}+1}\right)^{2(r-m)}\right) k^m (y_k \cos(kt + \pi m/2) + y_{k+\widehat{s}} \sin(kt + \pi m/2)) \end{aligned}$$

is optimal.

3. Criterion of existence of a linear method which is optimal for adaptive and nonadaptive recovery

Let $y \in \mathrm{Pr}_{\mathcal{Y}}\, G_\Phi$. The value

$$r(T, \mathcal{F}, \Phi, y) = \inf_{z \in Z} \sup_{x \in G_\Phi^{-1}(y)} \|Tx - z\|_Z$$

is called the Chebyshev radius of the set $T(G_\Phi^{-1}(y))$. It is the radius of the minimal ball containing the given set. If there exists such $z(y) \in Z$ that

$$r(T, \mathcal{F}, \Phi, y) = \sup_{x \in G_\Phi^{-1}(y)} \|Tx - z(y)\|_Z,$$

then $z(y)$ is called the Chebyshev center of the set $T(G_\Phi^{-1}(y))$.

The value

$$R^a(T, \mathcal{F}) = \inf_{\Phi\colon \mathcal{Y} \to \mathcal{P}} \sup_{y \in \mathrm{Pr}_{\mathcal{Y}}\, G_\Phi} r(T, \mathcal{F}, \Phi, y)$$

we call the adaptive radius of information in the problem of optimal adaptive recovery.

Lemma 2. *The following equality*

$$E^a(T, \mathcal{F}) = R^a(T, \mathcal{F})$$

holds.

Proof. Let $\varphi\colon \mathcal{Y} \to Z$ be an arbitrary method of recovery. Then for all $y \in \mathrm{Pr}_{\mathcal{Y}}\, G_\Phi$

$$r(T, \mathcal{F}, \Phi, y) \le \sup_{x \in G_\Phi^{-1}(y)} \|Tx - \varphi(y)\|_Z \le e^a(T, \mathcal{F}, \Phi, \varphi).$$

Taking the upper bound in the left-hand side over all $y \in \mathrm{Pr}_{\mathcal{Y}}\, G_\Phi$ and then taking the lower bound over all methods, we obtain

$$\sup_{y \in \mathrm{Pr}_{\mathcal{Y}}\, G_\Phi} r(T, \mathcal{F}, \Phi, y) \le \inf_{\varphi\colon \mathcal{Y} \to Z} e^a(T, \mathcal{F}, \Phi, \varphi).$$

Hence

$$R^a(T, \mathcal{F}) \le E^a(T, \mathcal{F}). \tag{7}$$

Let us prove the opposite inequality. Let $\varepsilon > 0$. For any $y \in \mathrm{Pr}_{\mathcal{Y}}\, G_\Phi$ there exists a $z_\varepsilon(y) \in Z$ such that

$$\sup_{x \in G_\Phi^{-1}(y)} \|Tx - z_\varepsilon(y)\|_Z \le r(T, \mathcal{F}, \Phi, y) + \varepsilon.$$

We define the method $\varphi_\varepsilon\colon \mathcal{Y} \to Z$ as follows

$$\varphi_\varepsilon(y) = \begin{cases} z_\varepsilon(y), & y \in \mathrm{Pr}_{\mathcal{Y}}\, G_\Phi, \\ 0, & y \notin \mathrm{Pr}_{\mathcal{Y}}\, G_\Phi. \end{cases}$$

Then

$$e^a(T,\mathcal{F},\Phi,\varphi_\varepsilon) = \sup_{(x,y)\in G_\Phi} \|Tx-\varphi_\varepsilon(y)\|_Z$$
$$= \sup_{y\in \Pr_{\mathcal{Y}} G_\Phi} \sup_{x\in G_\Phi^{-1}(y)} \|Tx-\varphi_\varepsilon(y)\|_Z \le \sup_{y\in \Pr_{\mathcal{Y}} G_\Phi} (r(T,\mathcal{F},\Phi,y)+\varepsilon).$$

Consequently,

$$E^a(T,\mathcal{F}) \le R^a(T,\mathcal{F})+\varepsilon.$$

In view of arbitrariness of $\varepsilon>0$ we obtain

$$E^a(T,\mathcal{F}) \le R^a(T,\mathcal{F}).$$

Together with (7) this proves the statement of lemma. □

Denote by $\operatorname{bco} A$ the convex balanced hull of A. It consists of all elements of the form

$$x = \lambda_1 x_1 + \ldots + \lambda_m x_m, \quad x_j \in A,\ j=1,\ldots,m, \quad \sum_{j=1}^m |\lambda_j| \le 1.$$

Let X' be the space algebraically dual to X, that is the space of all linear functionals on X and $Tx = \langle x', x\rangle$, where $x' \in X'$. We define the mapping $\operatorname{bco} F_p$ by its graph as follows

$$\operatorname{gr} \operatorname{bco} F_p = \operatorname{bco} \operatorname{gr} F_p.$$

Denote by $\mathcal{F}^{bco}$ the set of all mappings $\operatorname{bco} F_p$.

Theorem 5. *For the existence of linear nonadaptive method with the optimal parameter $\widehat{p}\in\mathcal{P}$ which is optimal over all adaptive methods it is necessary and sufficient that*

$$R^a(x',\mathcal{F}) = R^a(x',\mathcal{F}^{bco}) \tag{8}$$

and

$$D(\widehat{p}) = \inf_{p\in\mathcal{P}} D(p). \tag{9}$$

Proof. Necessity. Let the method $\widehat{\varphi}(y) = \langle y', y\rangle$ and the parameter $\widehat{p}\in\mathcal{P}$ such that

$$e(x',\mathcal{F},\widehat{p},\widehat{\varphi}) = E^a(x',\mathcal{F}).$$

Assume that $(x,y) \in \operatorname{bco}\operatorname{gr} F_{\widehat{p}}$. Then

$$(x,y) = \sum_{j=1}^m \lambda_j (x_j, y_j),$$

where $(x_j, y_j) \in \operatorname{gr} F_{\widehat{p}}$, $j = 1, \dots, m$, and $\sum_{j=1}^{m} |\lambda_j| \le 1$. Consequently,

$$|\langle x', x\rangle - \langle y', y\rangle| = \left|\sum_{j=1}^{m} \lambda_j(\langle x', x_j\rangle - \langle y', y_j\rangle)\right| \\ \le \max_{1\le j\le m} |\langle x', x_j\rangle - \langle y', y_j\rangle| \le e(x', \mathcal{F}, \widehat{p}, \widehat{\varphi}).$$

This implies that

$$e(x', \operatorname{bco}\mathcal{F}, \widehat{p}, \widehat{\varphi}) \le e(x', \mathcal{F}, \widehat{p}, \widehat{\varphi}).$$

Since

$$\operatorname{gr} F_p \subset \operatorname{gr}\operatorname{bco} F_p, \tag{10}$$

we have

$$e(x', \operatorname{bco}\mathcal{F}, \widehat{p}, \widehat{\varphi}) \ge e(x', \mathcal{F}, \widehat{p}, \widehat{\varphi}).$$

Thus we obtain

$$e(x', \operatorname{bco}\mathcal{F}, \widehat{p}, \widehat{\varphi}) = e(x', \mathcal{F}, \widehat{p}, \widehat{\varphi}). \tag{11}$$

Hence,

$$E(x', \operatorname{bco}\mathcal{F}) \le e(x', \operatorname{bco}\mathcal{F}, \widehat{p}, \widehat{\varphi}) = e(x', \mathcal{F}, \widehat{p}, \widehat{\varphi}) = E^a(x', \mathcal{F}).$$

In view of (4) we obtain

$$E^a(x', \operatorname{bco}\mathcal{F}) \le E^a(x', \mathcal{F}). \tag{12}$$

It follows from (10) that $E^a(x', \operatorname{bco}\mathcal{F}) \ge E^a(x', \mathcal{F})$. This with (12) implies that

$$E^a(x', \operatorname{bco}\mathcal{F}) = E^a(x', \mathcal{F}). \tag{13}$$

By virtue of Lemma 2 we obtain (8).

Now we prove (9). As it was noted above for linear functionals for any $p \in \mathcal{P}$ the following equality

$$e(x', \operatorname{bco}\mathcal{F}, p) = D(p) \tag{14}$$

holds. Taking into account (11) and (13) we have

$$D(\widehat{p}) = e(x', \operatorname{bco}\mathcal{F}, \widehat{p}) \le e(x', \operatorname{bco}\mathcal{F}, \widehat{p}, \widehat{\varphi}) = e(x', \mathcal{F}, \widehat{p}, \widehat{\varphi}) \\ = E^a(x', \mathcal{F}) = E^a(x', \operatorname{bco}\mathcal{F}) \le E(x', \operatorname{bco}\mathcal{F}) = \inf_{p\in\mathcal{P}} D(p).$$

Sufficiency. Assume that (8) and (9) hold. In view of the fact that the set $\operatorname{bco}\operatorname{gr} F_p$ is convex and balanced, it follows from [4] that for any $p \in \mathcal{P}$ there exists a $\widehat{\varphi}_p(y) = \langle y'_p, y\rangle$ such that

$$e(x', \operatorname{bco}\mathcal{F}, p) = e(x', \operatorname{bco}\mathcal{F}, p, \widehat{\varphi}_p).$$

Set $\widehat{\varphi} = \widehat{\varphi}_{\widehat{p}}$. Then from (14) and (9) we obtain

$$e(x', \operatorname{bco}\mathcal{F}, \widehat{p}, \widehat{\varphi}) = E(x', \operatorname{bco}\mathcal{F}).$$

Moreover, it follows from (14) and Theorem 3 that

$$E(x', \operatorname{bco} \mathcal{F}) = E^a(x', \operatorname{bco} \mathcal{F}).$$

Therefore, taking into account (11), we obtain

$$\begin{aligned} e(x', \mathcal{F}, \widehat{p}, \widehat{\varphi}) = e(x', \operatorname{bco} \mathcal{F}, \widehat{p}, \widehat{\varphi}) &= E^a(x', \operatorname{bco} \mathcal{F}) \\ &= R^a(x', \operatorname{bco} \mathcal{F}) = R^a(x', \mathcal{F}) = E^a(x', \mathcal{F}). \end{aligned}$$

□

References

[1] N.S. Bakhvalov, On the optimality of linear methods for operator approximation in convex classes of functions, USSR Comput. Math. Math. Phys. 11 (1971) 244–249.

[2] S. Gal, C.A. Micchelli, Optimal sequential and non-sequential procedures for evaluating a functional, Appl. Anal. 10 (2) (1980) 105–120.

[3] M.A. Kon, E. Novak, On the adaptive and continuous information problems, J. Complexity 5 (1989) 345–362.

[4] G.G. Magaril-Il'yaev, K.Yu. Osipenko, Optimal recovery of functionals based on inaccurate data, Math. Notes 50 (6) (1991) 1274–1279.

[5] G.G. Magaril-Il'yaev, K.Yu. Osipenko, On best harmonic synthesis of periodic functions, J. Math. Sci. 209 (1) (2015) 115–129.

[6] S.M. Nikolskii, Quadrature Formulas, Nauka, Moscow, 1988 (in Russian).

[7] L. Plaskota, Noisy Information and Computational Complexity, Cambridge University Press, 1996.

[8] J.F. Traub, H. Woźniakowski, A General Theory of Optimal Algorithms, New York, Academic Press, 1980.