КОНСТРУКТИВНАЯ ТЕОРИЯ ФУНКЦИЙ'81
София, 1983, с. 194—196

CONSTRUCTIVE FUNCTION THEORY'81
Sofia, 1983, p. 194—196

A REMARK ON OPTIMAL RECOVERY

L. Adamczyk

Summary. In this note we formulate the problem of optimal recovery as a point-to-set map approximation problem by single-valued function. The error of the best approximation is the intrinsic error of the estimation. Among all possible methods of estimating we are interested in those having the continuity properties.

1. Notation and Background. The setting of this paper is adopted from Micchelli and Rivlin [4], which also contains a survey of most of the results in this area. Let K be a subset of a linear space X and Y, Z two normed linear spaces. Consider a linear or nonlinear operator U such that $U:K\to Z$. We want to estimate Ux from the incomplete information about x. A priori information on x is available in the statement that $x\in K$. At our disposal there is also the approximate value y of an information operator, not necessarily linear, I from D into Y, where $K\subset D\subset X$, that is $\|Ix-y\|\leqq\varepsilon$. An estimation algorithm A is then a, not necessarily linear, transformation from $Q:=IK+\varepsilon S$ $S:=\{y\in Y: \|y\|\leqq 1\}$, into Z, $\varepsilon\geqq 0$ — a preassigned tolerance.

The algorithm A produces an error $E_A(K,\varepsilon):=\sup\{\|Ux-Ay\|: x\in K, \|Ix-y\|\leqq\varepsilon\}$. We call $E(K,\varepsilon):=\inf_A E_A(K,\varepsilon)$ the error in the recovery problem $X, Y, Z, I, U, \varepsilon$ and K, and any algorithm A_0, satisfying

$$E_{A_0}(K,\varepsilon)=E(K,\varepsilon) \tag{1.1}$$

is called an optimal recovery algorithm.

For most problems the information operator I is not one-to-one. Thus there may exist many different elements $x\in K$ with the same information Let $x\in K$. Let $V(y):=\{v\in K: Ix=Iv, \|Ix-y\|\leqq\varepsilon\}$ be the pre-image set of all Ix such that $\|Ix-y\|\leqq\varepsilon$. Note that $V(y)$ is not empty since $y\in Q$. Furthermore let $U(y):=\{Ux: x\in V(y)\}$ be the set of all Ux which share the information $\|Ix-y\|\leqq\varepsilon$. In this way at each recovery problem $X, Y, Z, I, U, \varepsilon$ and K the following multifunction

$$F: Q\to 2^Z, \text{ where } F(y):=U(V(y)) \tag{1.2}$$

is uniquely determinated. As we shall see below, each optimal algorithm defined by (1.1) is a best approximation function for the point-to-set map F in the sense given in part 2.

2. Point-to-Set Map Approximation Problem. Recall [1], that for a set $A\subset Z$ rad $(A):=\inf_{x\in Z}\sup_{x\in A}\|x-v\|$ is called the radius of A.

Roughly speaking, rad (A) is the minimal radius of a "ball" which contains A If there exists c, $c \in Z$, such that $\sup_{y \in A} \|c-y\| = \text{rad}(A)$, then c is a center of A or, in other terminology [2], Chebyshev center of A. Note that c can be an element outside A and need not be unique [2].

Suppose F is a map from Q into subsets of Z. For any function $f: Q \to Z$ define the distance of an f from F by the relation

$$d(f, F) := \sup_{x \in Q} \sup_{y \in F(y)} \|f(x) - y\|. \tag{2.1}$$

An f_0 is called the best approximation of a set-valued function F if $d(f_0, F) := \inf_f d(f, F)$. Put

$$r(F) := \sup_{x \in Q} \text{rad}(F(x)). \tag{2.2}$$

For an arbitrary F, rad $(F(x))$ for some x or $r(F)$ may be infinite.

Theorem 2.1. *An algorithm A_0 is an optimal one iff A_0 is a best approximation of a set-valued function F given by* (1.2). *Furthermore, equality*

$$r(F) = E(K, \varepsilon) \tag{2.3}$$

holds.

Proof. Let A be any transformation from Q into Z. We have the inequality $\sup_{x \in F(y)} \|Ay - x\| \geq \text{rad}(F(y))$ for each $y \in Q$. Therefore by (2.1) and (2.2) we get $d(A, F) \geq r(F)$ for each A. Hence, also

$$E(K, \varepsilon) = \inf_A d(A, F) \geq r(F). \tag{2.4}$$

On the other hand, let $\delta \geq 0$ be an arbitrary number. Define an algorithm A_δ as follows.

Let $A\,y := z_\delta$, where $\|z_\delta - x\| \leq \text{rad}(F(y)) + \delta$ for all $x \in F(y)$. Thus z_δ is almost a center of $F(y)$. Then $E(K, \varepsilon) \leq E_{A_\delta}(K, \varepsilon) = \sup_{y \in Q} \sup_{z \in F(y)} \|Ay - z\| \leq \sup_{y \in Q} \text{rad}(F(y)) + \delta = r(F) + \delta$. Since δ is arbitrary, $E(K, \varepsilon) \leq r(F)$ and by (2.4) the equality (2.3) holds. For the rest, we have

$$E(K, \varepsilon) = \inf_A E_A(K, \varepsilon) = \inf_A \sup\{\|Ux - Ay\| : x \in K, \|Ix - y\| \leq \varepsilon\}$$
$$= \inf_A \sup_{y \in Q} \sup_{z \in F(y)} \|z - Ay\| = \inf_A d(A, F) = r(F).$$

So for an optimal algorithm A_0 there is $E(K, \varepsilon) = E_{A0}(K, \varepsilon) = r(F) = d(A_0, F)$. See Micchelli and Rivlin [4], where a similar result is established for I and U linear. Theorem 2.1 motivates the using a center $c(y)$ of $F(y)$, if it exists, as an approximation for all elements of $F(y)$.

Definition 2.2. An algorithm $A^c: Q \to Z$ is a central algorithm iff for each $y \in Q$, $A^c(y)$ is a center of $F(y)$.

Theorem 2.3. *Any central algorithm is an optimal algorithm, i. e.* $E_A c(K, \varepsilon) = d(A^c, F) = r(F)$.

Proof. Note that $E_A c(K, \varepsilon) = \sup_{y \in Q} \sup_{z \in F(y)} \|z - A^c y\| = \sup_{y \in Q} \text{rad}(F(y)) = r(F)$.

Theorem 2.4. *If the normed space Z is reflexive and strictly convex and $r(F)$ is finite, there exists one and only one central algorithm for F.*

Following Traub and Woźniakowski [8] we introduce

Definition 2.5. An algorithm $A^I: Q \to Z$ is an interpolatory algorithm iff in recovery problem F: $\varepsilon = 0$ and $A^I y \in F(y)$ for $y \in Q$, i. e. for each observed $y \in Q$, there is $x \in K$ such that $Ix = y$ and $Ux = A^I y$. As it was noted in [8], an interpolatory algorithm may turn out to be an optimal algorithm. In fact, we have the following

Theorem 2.6. *Suppose that for recovery problem F we have: Z is Hilbert space and all $F(y)$, for $y \in Q$, are closed convex bounded subsets of Z. There exists a unique central algorithm which is the interpolatory algorithm.*

3. Continuity Properties of Optimal Algorithms as a Function of Data. Among all possible methods of estimating an operator U from inaccurate data we are interested in those having the continuity properties.

Theorem 3.1. *For a given recovery problem F, let Z be a uniformly convex space. Then, if the map F is closed-valued and upper semi-continuous, there exists an $f_0 \in C(Q,Z)$ such that $d(f_0, F) = r(F)$.*

In other words, f_0 is the continuous optimal algorithm as a function of data.

In the next part we analyse continuity properties of central algorithms.

Definition 3.2 (cf. Singer [7]). A set-valued function F mapping a topological space Q into 2^Z is said to be upper Hausdorff semicontinuous H. u. s. c. if for every $y_0 \in Q$ and every $\varepsilon \geq 0$ there is a neighbourhood Q_0 of y_0 such that for every $y \in Q_0$ we have $\sup_{x \in F(y)} \operatorname{dist}(x, F(y_0)) \leq \varepsilon$.

Definition 3.3 (cf. Singer [6]). Normed linear space Z is said to have Efimov-Steczkin (ES-) property if every weakly-closed set in Z is approximatively compact. If in addition Z is strictly convex it is said to be E-space (cf. Holmes [2]).

Let $\check{C}(B)$ denote the set of all centers of set B, and let $\check{C}$ also denote, in general, the point-to-set map $\check{C}: B \to \check{C}(B)$.

Theorem 3.4. *Let Z be an ES-space. Suppose that the multifunction $F: Q \to 2^Z$ is compact valued and Hausdorff continuous. Then the point-to-set map $\check{C}F: Q \to 2^Z$, where $\check{C}F(y) := \check{C}(F(y))$, is also compact valued and H.u.s.c.*

Corollary 3.5. Let Z be an E-space and a recovery problem F be compact valued and Hausdorff continuous. Then unique central algorithm is continuous.

In particular this result holds in a uniformly convex space, since every uniformly convex space is an E-space (of [2]).

However, in a uniformly convex space we have much stronger result.

Theorem 3.6. *Let Z be a uniformly convex space. Suppose that $h := h(A, B)$ is a Hausdorff distance between sets A and B. Let $a := \check{C}(A)$ and $b := \check{C}(B)$. Let $\varepsilon \geq 0$ be arbitrary and $\delta = \delta(\varepsilon)$ be modulus of convexity. If $h \leq \delta \cdot r/(2(1-\delta))$ then $\| a - b \| \leq \varepsilon(r + 2h)$, where $r =$* rad (A) *or* rad (B).

In this way a uniformly continuous central algorithm is possible for recovery problems with space Z being uniformly convex.

REFERENCES

1. А. Гаркави. О наилучшей сети и наилучшем сечении множества в нормированном пространстве. *Изв. Акад. наук СССР*, **26**, 1962, 87-106.
2. R. B. Holmes. A course on optimization and best approximation. *Lect. Notes in Math.*, **257**.
3. A. A. Melkman, Ch. A. Micchelli. Optimal estimation of linear operators in Hilbert spaces from inaccurate data. *SIAM J. Numer. Anal.*, **16**, 1979, No 1.
4. Ch. A. Micchelli, T. J. Rivlin. A Survey of Optimal Recovery. Optimal Estimation in Approximation Theory. New York, 1977, 1-53.
5. Ch. A. Micchelli, T. J. Rivlin. Optimal recovery of best approximations. *Resultate der Math.*, **3**, 1980, 25-32.
6. I. Singer. Some remarks on approximative compactness. *Rev. Roum. Math. Pures Appl.*, **9**, 1964, No 1, 167-177.
7. I. Singer. The theory of best approximation and functional analysis. — In: Reg. Conf. Ser. Appl. Math., 13, SIAM. Ph., 1974.
8. J. F. Traub, H. Woźniakowski. General theory of optimal error algorithms and analytic complexity. Part A. — In: General Information Model. Dept. Comput. Sci. Cornegie-Mellon Univ. September 1977.

Oscar Lange Academy of Economics
Institute of Cybernetics
53-345 Wrocław Poland

Received on June 3, 1981