JOURNAL OF APPROXIMATION THEORY 45, 201–218 (1985)

Lipschitz Condition in Minimum Norm Problems on Bounded Functions*

VASANT A. UBHAYA

Department of Computer Science and Operations Research, North Dakota State University, Fargo, North Dakota 58105, U.S.A.

Communicated by Oved Shisha

Received April 14, 1983; revised March 25, 1985

Consider the Banach space of bounded functions with uniform norm. Given an element f and a closed convex set in this space, the minimum norm problem is to find an element in the convex set nearest to f. Such a nearest point is not unique in general. For each f in the space, is it possible to select a nearest element f' so that the selection operator $f \to f'$ satisfies a Lipschitz condition with some constant C? If so, does there exist an operator for which C is minimum? This article determines the required Lipschitzian selection operators with smallest possible constants for the minimum norm problem in three cases of special interest.

1. INTRODUCTION

It is a well-known fact that in a Hilbert space, the projection operator, mapping the space onto a closed linear manifold, satisfies a Lipschitz condition with constant unity. Indeed, if f' and h' are respective projections of two elements f and h of the Hilbert space onto the manifold, then $\|f' - h'\| \leqslant \|f - h\|$, where $\|\cdot\|$ is the Hilbert norm. More generally, this condition also holds when f' is the unique element closest to f in a closed convex subset of a Hilbert space. Now consider the minimum norm problem of finding a nearest element from a closed convex set in the Banach space of bounded functions with uniform norm. This article establishes the Lipschitz condition with appropriate constants for operators in this problem. Given an element f in this space, the set of all elements in the convex set nearest to f is itself convex and in general, this set is not a singleton as it is in a Hilbert space. Questions that arise naturally in this

* An abstract of this paper appears in *Abstracts Amer. Math. Soc.* 4 (1983), 396. This paper was presented by the author at the Joint National ORSA/TIMS Meeting in Chicago in April 1983.

0021-9045/85 $3.00

case are the following. For each f, is it possible to select an element f' from this convex set of nearest elements so that the non-linear selection operator $f \to f'$ satisfies the Lipschitz condition with some constant C, i.e., $\|f' - h'\| \leqslant C \|f - h\|$ holds for all pairs of elements f and h? If so, does there exist an optimal selection operator for which C is minimum? In this article, such optimal Lipschitzian selection operators are constructed for three problems of special interest, viz., the greatest convex minorant, approximation by convex functions and generalized isotone optimization. It is interesting to note that Hilbert-like properties apply to the nonlinear operators constructed in these problems although the space under consideration is not Hilbert.

We now introduce some notation and elaborate on the problem. Let S be any set and B denote the Banach space of all bounded real functions on S with the uniform norm,

$$\|f\| = \sup\{|f(s)| : s \in S\}, \qquad f \in B. \tag{1.1}$$

Given an element f in B, let K_f, possibly dependent on f, denote a non-empty closed convex subset of B. If $\Delta(f)$ denotes the infimum of $\|f - k\|$ over all k in K_f, the set G_f of all nearest elements g satisfying

$$\Delta(f) = \|f - g\| = \inf\{\|f - k\| : k \in K_f\}, \tag{1.2}$$

is a closed convex subset of B, as may be easily verified. A selection operator T' is a nonlinear operator which maps each f in B to an f' in G_f. We wish to determine an optimal selection operator $T: f \to f'$, and a (least) number C, if these exist, so that

$$\|T(f) - T(h)\| \leqslant C \|f - h\| \qquad \text{for all } f,\ h \in B, \tag{1.3}$$

where T is such that C is the smallest number for all possible choices of the selection operator T'. That is

$$C = \inf \sup\{\|T'(f) - T'(h)\| / \|f - h\| : f, h \in B, f \neq h\}, \tag{1.4}$$

where the infimum is taken over all T' and is attained at T. In addition, we examine the validity of

$$|\Delta(f) - \Delta(h)| \leqslant D \|f - h\| \qquad \text{for all } f,\ h \in B, \tag{1.5}$$

where D is the smallest possible number satisfying (1.5) or

$$D = \sup\{|\Delta(f) - \Delta(h)| / \|f - h\| : f, h \in B, f \neq h\}. \tag{1.6}$$

Note that $\Delta(f)$ is independent of T. In each of the three problems considered in this article, we first determine an operator T with its associated constant C. We then establish the optimality of T by showing that a lower value of C does not exist for any T'. In Section 2 we show that the constant, say $C(T')$, associated with any Lipschitzian selection operator T' is a convex function of T' over some convex set of operators. Thus, $C = C(T)$ minimizes this convex function. In Section 5, we give examples of selection operators with different values of C. The issue of uniqueness of T is not analyzed in this article; it will be considered elsewhere.

The motivation for considering this problem comes from a similar result in a Hilbert space mentioned earlier (e.g., [7, p. 100]) and this author's earlier result [18, II, p. 320] for the problem of isotone optimization. In Section 3, we consider the problem of finding a convex function nearest to a given f but not exceeding it at any point. We show that an optimal T maps f to its greatest convex minorant, the maximal optimal solution to the problem, with $C = 1$ and $D = 2$. In Section 4, we analyze the problem of finding a convex function nearest to f and show that an optimal T maps f to the maximal optimal solution of the problem with $C = 2$ and $D = 1$. It is shown in [19] that this maximal optimal solution is the greatest convex minorant of f shifted upward through a certain distance. This observation, together with an available algorithm for obtaining convex hulls [1, 2, 8, 9, 14], has led to a linear time algorithm ($O(n)$) for computing this solution on a set of n points in an interval. Other linear time algorithms based on linear programming approaches appear in [21]. In Section 5, an optimal T with $C = 1$ is constructed for the problem of generalized isotone optimization [20] when the weight function is identically equal to unity. Here, S is a partially ordered set and the convex set under consideration is a closed convex cone determined by isotonicity and nonnegativity conditions on functions. It is shown in [20] that any optimal solution to the problem is "enclosed" between minimal and maximal optimal solutions. The operator T maps f to the mean of these two solutions.

Minimum norm problems arise as curve fitting or estimation problems when the initial data points $f(t)$ based on experimental observations display certain random variations and need be estimated by an element from a convex set K_f. We write $f(t) = \mu(t) + \eta(t)$, where μ is in K_f and η represents a random disturbance or noise. The actual values of μ are not known. We estimate μ by f' which is in K_f, is nearest to f and has additional properties. For example, in economics, assumptions of concavity or convexity are often made regarding various functions such as utility, marginal utility, production, etc. [10, 11]. If $\mu(t)$ is such a convex function representing a particular entity as a function of t, we obtain its convex estimate $f'(t)$ on the basis of the actual observations $f(t)$ of the entity. A special case of the problem of generalized isotone optimization discussed in Section 5 arises, for example, when it becomes necessary to estimate the failure rate of a system under the assumption that it is nonincreasing. This assumption applies during the "debugging" period of the system when the defects of the system are gradually being eliminated [18, 20]. Inequality (1.3) involving an optimal Lipschitzian selection operator $f \to f'$ signifies the minimum possible sensitivity of a nearest element f' to changes in f. Consequently, f' is the most desirable estimate of f. Additional references on analysis of similar problems are [4, 5, 13, 17, 22, 23]. A survey of constrained approximation problems on the space of continuous functions in which the approximants mainly form a subset of a Haar subspace or are rational functions, appears in [3]. Although, most problems considered in this survey have a different structure from those analyzed in this article, the underlying concept—approximation from convex subsets—is the same.

2. A General Observation

In this section we consider problem (1.2) in a general setting and establish convexity of the constant associated with Lipschitzian selection operators. We do not attempt any complete treatment of the general case.

For any Lipschitzian selection operator T', let $C(T')$ be the smallest number satisfying

$$\|T'(f) - T'(h)\| \leqslant C(T')\, \|f-h\| \qquad \text{for all } f,\ h \in B.$$

Let X be the vector sace of all operators with domain and range B. Let X' be the subset of X consisting of all Lipschitzian selection operators T'.

Proposition 2.1. *X' is a convex subset of X and $C(T')$ is a convex function of T' over X'.*

Proof. Let $T_1, T_2 \in X'$. Let also $0 \leqslant \lambda \leqslant 1$ and $T_3 = \lambda T_1 + (1-\lambda)\, T_2$. Since $T_1(f)$, $T_2(f) \in G_f$ and G_f is convex, we conclude that $T_3(f) \in G_f$. Now for all $f, h \in B$ we have

$$\begin{aligned}\|T_3(f) - T_3(h)\| &\leqslant \lambda\, \|T_1(f) - T_1(h)\| + (1-\lambda)\, \|T_2(f) - T_2(h)\| \\ &\leqslant (\lambda C(T_1) + (1-\lambda)\, C(T_2))\, \|f-h\|.\end{aligned}$$

Hence, $T_3 \in X'$ and X' is convex. Also

$$C(T_3) \leqslant \lambda C(T_1) + (1-\lambda)\, C(T_2),$$

which establishes the convexity of $C(T')$. The proof is now complete.

3. The Greatest Convex Minorant

In this section we consider the problem of finding a convex function nearest to a given function f but not exceeding it at any point. The greatest convex minorant of f is the maximal optimal solution to this problem. We show that the operator mapping f to its greatest convex minorant is an optimal Lipschitzian selection operator with $C=1$ and $D=2$.

Let $S=I=[a,b]$, a closed real interval and B the set of all bounded functions on I with uniform norm (1.1). A function k in B is said to be convex if $k(\lambda s+(1-\lambda)t)\leqslant\lambda k(s)+(1-\lambda)k(t)$ for all s, t in I and all $0\leqslant\lambda\leqslant 1$ [15]. For each f in B, we let

$$K_f=\{k: k \text{ is convex and } k(s)\leqslant f(s) \text{ for all } s\in I\}, \tag{3.1}$$

and consider problem (1.2). Clearly, K_f is a nonempty closed convex subset. For notational convenience later, we replace $\Delta(f)$ of (1.2) by $\bar{\Delta}(f)$. Thus (1.2) becomes

$$\bar{\Delta}(f)=\|f-g\|=\inf\{\|f-k\|: k\in K_f\}. \tag{3.2}$$

We observe that a convex function is continuous in the interior of I [15].

We define the greatest convex minorant $\bar{f}$ of f to be the largest convex function which does not exceed f at any point in I. Specifically,

$$\bar{f}(s)=\sup\{k(s): k\in K_f\} \qquad \text{all } s\in I.$$

Since the pointwise supremum of a set of convex functions is convex [15], it follows that $\bar{f}$ is convex with $\bar{f}\leqslant f$. It is easy to show that $\bar{f}$ minimizes $\|f-k\|$ for k in K_f and thus $\bar{\Delta}(f)=\|f-\bar{f}\|$. In addition, $\bar{f}\geqslant g$ for all g in G_f and hence $\bar{f}$ is the maximal optimal solution to the problem. The following example, to be used later, illustrates that a minimizer g in (3.2) is not unique in general. Define f_0 on $[0,1]$ by

$$\begin{aligned} f_0(s) &= -1, \qquad s=0,\\ &= \;\;\,1, \qquad 0<s\leqslant 1. \end{aligned} \tag{3.3}$$

It is easy to verify that $\bar{f}_0(s)=2s-1$, $0\leqslant s\leqslant 1$ and when $f=f_0$,

$$G_f=\{g: g \text{ is convex and } -1\leqslant g(s)\leqslant\bar{f}_0(s) \text{ for all } 0\leqslant s\leqslant 1\}.$$

We now state our main result for the mapping $f\to\bar{f}$.

THEOREM 3.1. *Define $T: B\to B$ by $T(f)=\bar{f}$ where $\bar{f}$ is the greatest convex minorant of f in B. Then*

$$\|T(f)-T(h)\|\leqslant\|f-h\| \qquad \textit{for all } f, h\in B, \tag{3.4}$$

and

$$|\bar{\Delta}(f)-\bar{\Delta}(h)| \leqslant 2\,\|f-h\| \qquad \textit{for all } f, h \in B. \tag{3.5}$$

T is an optimal Lipschitzian selection operator with $C=1$ *and* $D=2$ *in* (1.3) *and* (1.5), *respectively.*

In order to prove Theorem 3.1, we establish some preliminary results. For any f in B and t in I, we let

$$L(f, t)=\bigcup \{(u, v)\colon a \leqslant u<v \leqslant b,\ t \in(u, v),\ f \text{ is linear on } (u, v)\}. \tag{3.6}$$

We define the linear set $L(f)$ of f by

$$L(f)=\bigcup \{L(f, t)\colon t \in I\}. \tag{3.7}$$

LEMMA 3.1. *$L(f, t)$ is either empty or is some nonempty open interval (c_t, d_t) on which f is linear.*

Proof. Assume that $L(f, t)$ is nonempty. Define

$$c=c_t=\inf\{u\colon a \leqslant u<v \leqslant b,\ t \in(u, v),\ f \text{ is linear on } (u, v)\},$$

$$d=d_t=\sup\{v\colon a \leqslant u<v \leqslant b,\ t \in(u, v),\ f \text{ is linear on } (u, v)\}.$$

Clearly $c<d$. Since (c, d) contains each (u, v) in (3.6), we have $(c, d) \supset L(f, t)$. Let $x, y \in(c, d)$ with $x<y$. Let also $0<\lambda<1$ and $z=\lambda x+(1-\lambda)\, y$. We show that $L(f, t)=(c, d)$ and $f(z)=\lambda f(x)+(1-\lambda)\, f(y)$, i.e., f is linear on (c, d). By the definition of c and d, there exist intervals (u, v) and (r, s) containing t such that f is linear on each of (u, v), (r, s), and $u<x<y<s$. It is easy to show that f is linear on (u, s) and thus (u, s) is one of the open intervals in (3.6). It follows that $x, y \in L(f, t)$ and $L(f, t)=(c, d)$. Again by linearity of f on (u, s) we have $f(z)=\lambda f(x)+(1-\lambda)\, f(y)$. Thus f is linear on (c, d). The proof is now complete.

LEMMA 3.2. *If $L(f, t) \cap L(f, s)$ is nonempty then $L(f, t)=L(f, s)$.*

Proof. Let $u \in L(f, t) \cap L(f, s)$. By Lemma 3.1, $L(f, t)=(c_t, d_t)$ and $L(f, u)=(c_u, d_u)$. Since $u \in(c_t, d_t)$ and f is linear on (c_t, d_t), we conclude from the definition of $L(f, u)$ that $(c_t, d_t) \subset L(f, u)$. Hence $t \in(c_u, d_u)$ and $(c_u, d_u) \subset L(f, t)$. Thus $L(f, t)=L(f, u)$. Similarly $L(f, s)=L(f, u)$. The proof is now complete.

PROPOSITION 3.1. *If $L(f)$ is nonempty for some f in B, then*

$$L(f)=\bigcup_n (c_n, d_n)$$

where the union is over a finite or countably infinite set of disjoint open intervals contained in I so that f is linear on each of them.

Proof. By Lemma 3.1, we have

$$L(f) = \bigcup \{(c_t, d_t): t \in I\}.$$

By Lemma 3.2, any two nonidentical intervals (c_t, d_t) are disjoint. Since each interval includes a distinct rational number, the countability of intervals follows and the required result is established.

PROPOSITION 3.2. *If $\bar{f}$ is the greatest convex minorant of f in B, then $\bar{f}(a) = f(a)$ and $\bar{f}(b) = f(b)$. If $t \in I - L(\bar{f})$ then $\bar{f}(t) = f(t)$ or there exists a sequence $\langle t_n \rangle$ of points in I with $t_n \neq t$ for each n such that $t_n \to t$ and $f(t_n) \to \bar{f}(t)$.*

Proof. We first show that $\bar{f}(s) = f(s)$ if $s = a, b$. Define a function p in B by

$$\begin{aligned} p(s) &= f(s) \qquad \text{if} \quad s = a, b, \\ &= \bar{f}(s) \qquad \text{if} \quad s \in (a, b). \end{aligned}$$

Since $f \geqslant \bar{f}$ we see that $p \geqslant \bar{f}$. Clearly, p is convex and hence by the definition of $\bar{f}$, we have $f \geqslant \bar{f} \geqslant p$. Consequently, $\bar{f} = p$ and $\bar{f}(s) = f(s)$ for $s = a, b$.

Now we prove the remaining part. Let $t \in I - L(\bar{f})$ and $t \neq a$ or b. Assume that there exists an $\varepsilon > 0$ and an open interval (u, v) with $t \in (u, v)$ such that $f(s) - \bar{f}(t) > \varepsilon$ for all s in (u, v). We shall reach a contradiction. By continuity of $\bar{f}$ on (a, b), there exists an open interval $(x, y) \subset (u, v)$ such that $t \in (x, y)$ and

$$|\lambda \bar{f}(x) + (1 - \lambda)\bar{f}(y) - \bar{f}(t)| < \varepsilon \qquad \text{for all } 0 < \lambda < 1. \tag{3.8}$$

Combining this with the hypothesis $f(s) - \bar{f}(t) > \varepsilon$ for s in (u, v) we find that

$$f(s) - (\lambda \bar{f}(x) + (1 - \lambda)\bar{f}(y)) > 0 \qquad \text{for all } s \in (x, y), \text{ all } 0 < \lambda < 1. \tag{3.9}$$

Now define a convex function q on I by

$$\begin{aligned} q(s) &= \left(\frac{y - s}{y - x}\right)\bar{f}(x) + \left(\frac{s - x}{y - x}\right)\bar{f}(y), \qquad && s \in (x, y), \\ &= \bar{f}(s), && s \in I - (x, y). \end{aligned} \tag{3.10}$$

Note that q is linear on (x, y). By (3.10) and the convexity of $\bar{f}$, we have $q \geqslant \bar{f}$. Again, (3.9) and the fact that $f \geqslant \bar{f}$ show that $f \geqslant q$. Hence $q = \bar{f}$. Now since $t \notin L(f)$, $\bar{f}$ is not linear on (x, y). But q is linear on (x, y) which is a contradiction. This establishes the existence of the sequence $\langle t_n \rangle$ with properties as stated. The proof is now complete.

Before stating our next proposition, we observe that for a convex function k the right-hand (left-hand) limit $k(a+0)$ $(k(b-0))$ exists at $a(b)$ and $k(a) \geqslant k(a+0)$ $(k(b) \geqslant k(b-0))$ [15].

PROPOSITION 3.3. *Let $\bar{f}$ and $\bar{h}$ respectively be the greatest convex minorant of f and h in B. Then*

(i) $\bar{f}(a+0) = \min(f(a), \liminf\{f(s): s \to a\})$,
$\bar{f}(b-0) = \min(f(b), \liminf\{f(s): s \to b\})$.

(ii) $|\bar{f}(a+0) - \bar{h}(a+0)| \leqslant \|f-h\|$,
$|\bar{f}(b-0) - \bar{h}(b-0)| \leqslant \|f-h\|$.

Proof. We show the first equality in (i). The proof for the second equality is similar. By the convexity of $\bar{f}$ and Proposition 3.2, we have $\bar{f}(a+0) \leqslant \bar{f}(a) = f(a)$. Again since $\bar{f}(s) \leqslant f(s)$ for all s, by taking lim inf, we conclude that

$$\bar{f}(a+0) \leqslant \min(f(a), \liminf\{f(s): s \to a\}). \tag{3.11}$$

Now suppose that strict inequality holds in (3.11). We shall reach a contradiction. In this case, there exists a t in (a, b) such that

$$\sup\{\bar{f}(s): a < s \leqslant t\} < \inf\{f(s): a \leqslant s \leqslant t\} = \theta, \tag{3.12}$$

say. Now define,

$$\begin{aligned} f^0(s) &= f(a), && s = a, \\ &= \left(\frac{s-a}{t-a}\right)\bar{f}(t) + \left(\frac{t-s}{t-a}\right)\theta, && s \in (a, t), \\ &= \bar{f}(s), && s \in [t, b]. \end{aligned}$$

Clearly, f^0 is convex and by (3.12) $\bar{f} \leqslant f^0 \leqslant f$. Again by (3.12), $\bar{f}(s) < f^0(s)$ for all s in (a, t), which is a contradiction to the fact that $\bar{f}$ is the greatest convex minorant of f. Hence the first equality in (i) is established.

We now prove the first inequality in (ii). The proof for the second inequality is similar. We first show that

$$\bar{f}(a+0) - \bar{h}(a+0) \leqslant \|f-h\|. \tag{3.13}$$

Using (i), suppose first that $\bar{h}(a+0)=h(a)$. Then since $\bar{f}(a+0)\leqslant f(a)$, the inequality (3.13) immediately follows. Now suppose that

$$\bar{h}(a+0)=\lim\inf\{h(s): s\to a\}.$$

Then, given $\varepsilon>0$, there exists x in (a, b) satisfying simultaneously

$$|\bar{h}(a+0)-h(x)|<\varepsilon/2, \qquad \bar{f}(a+0)\leqslant f(x)+\varepsilon/2.$$

It follows that

$$\bar{f}(a+0)-\bar{h}(a+0)\leqslant f(x)-h(x)+\varepsilon\leqslant \| f-h\| +\varepsilon.$$

Thus, (3.13) is established. A symmetric argument completes the proof of the first inequality in (ii). The proof of the proposition is now complete.

Now we prove Theorem 3.1.

Proof of Theorem 3.1. We first establish (3.4). Suppose that $t\in I-L(\bar{h})$. Then by Proposition 3.2, $\bar{h}(t)=h(t)$ or $\bar{h}(t)=\text{limit } h(t_n)$ for some sequence $t_n\to t$, $t_n\neq t$. In the former case (and this case, by Proposition 3.2, includes $t=a$ or b), we have

$$\bar{f}(t)-\bar{h}(t)\leqslant f(t)-h(t)\leqslant \| f-h\|.$$

In the latter case assume $t\neq a$ and $t\neq b$. Since $\bar{f}\leqslant f$ and $\bar{f}$, $\bar{h}$ are continuous on (a, b), we have

$$\bar{f}(t)-\bar{h}(t)=\text{limit}(\bar{f}(t_n)-h(t_n))\leqslant \text{limit}(f(t_n)-h(t_n))\leqslant \| f-h\|.$$

Hence in either case we have

$$\bar{f}(t)-\bar{h}(t)\leqslant \| f-h\|. \tag{3.14}$$

Now suppose that $t\in L(\bar{h})$. Then by Proposition 3.1, $t\in(c, d)=(c_n, d_n)$ for some n and $\bar{h}$ is linear on (c, d). Let $t=\lambda c+(1-\lambda)\, d$ for some $0<\lambda<1$. Assume first that $c\neq a$ and $d\neq b$. Then

$$\bar{h}(t)=\lambda\bar{h}(c)+(1-\lambda)\,\bar{h}(d).$$

Again, by convexity of $\bar{f}$ we have

$$\bar{f}(t)\leqslant\lambda\bar{f}(c)+(1-\lambda)\,\bar{f}(d).$$

On subtraction, we find that

$$\bar{f}(t)-\bar{h}(t)\leqslant\lambda(\bar{f}(c)-\bar{h}(c))+(1-\lambda)(\bar{f}(d)-\bar{h}(d)). \tag{3.15}$$

Now $c, d \in I - L(\bar{h})$. Consequently, by the first part of the proof, (3.14) holds with $t = c$ and d. It follows from (3.15) that (3.14) holds for t. Now assume that $c = a$ and $d \neq b$. Then $t = \lambda a + (1 - \lambda)\, d$ and

$$\bar{h}(t) = \lambda \bar{h}(a + 0) + (1 - \lambda)\, \bar{h}(d).$$

Again, the function obtained from $\bar{f}$ by replacing $\bar{f}(a)$ by $\bar{f}(a + 0)$ is convex. Hence

$$\bar{f}(t) \leqslant \lambda \bar{f}(a + 0) + (1 - \lambda)\, \bar{f}(d).$$

On subtraction, we have

$$\bar{f}(t) - \bar{h}(t) \leqslant \lambda(\bar{f}(a + 0) - \bar{h}(a + 0)) + (1 - \lambda)(\bar{f}(d) - \bar{h}(d)).$$

By Proposition 3.3(ii) and arguing as before, we conclude that (3.14) holds for t. Other cases for which $d = b$ may be considered similarly. Thus (3.14) holds for all t in I. A symmetric argument shows that $\bar{h}(t) - \bar{f}(t) \leqslant \| f - h \|$ and this establishes (3.4).

To show (3.5), we let $\varepsilon > 0$. Then there exists s in I such that $\bar{\Delta}(f) \leqslant f(s) - \bar{f}(s) + \varepsilon$. Again $\bar{\Delta}(h) \geqslant h(s) - \bar{h}(s)$. On subtracting and using (3.4) we obtain

$$\begin{aligned} \bar{\Delta}(f) - \bar{\Delta}(h) &\leqslant f(s) - h(s) - (\bar{f}(s) - \bar{h}(s)) + \varepsilon \\ &\leqslant \| f - h \| + \| \bar{f} - \bar{h} \| + \varepsilon \leqslant 2\, \| f - h \| + \varepsilon. \end{aligned}$$

Thus $\bar{\Delta}(f) - \bar{\Delta}(h) \leqslant 2\, \| f - h \|$. A symmetric argument completes the proof of (3.5). (An alternative proof of (3.5) is given at the end of Sect. 4.)

We now show that T is optimal. We show that $C = 1$ is optimal in the sense of (1.4). Indeed, if f and h are two distinct convex functions then $G_f = \{f\}$ and $G_h = \{h\}$. Consequently, for any selection operator T', we must have $T'(f) = f$ and $T'(h) = h$. Hence (1.3) with $T = T'$ shows that $C \geqslant 1$. Since (3.4) shows that $C \leqslant 1$ the optimality of $C = 1$ and hence of T is established. We now show that $D = 2$ is optimal in the sense of (1.6). Consider f_0 defined by (3.3). Let h on $[0, 1]$ be identically zero. Then $\bar{\Delta}(f_0) = 2$, $\bar{\Delta}(h) = 0$, and $\| f_0 - h \| = 1$. We see from (1.5) that $D \geqslant 2$. But since (3.5) shows that $D \leqslant 2$, the optimality of $D = 2$ is established. The proof of the theorem is now complete.

We now make a remark. It is possible to establish Theorem 3.1 by a shorter proof similar to the one for the following result in fixed point theory: the mapping of closed bounded subsets onto their closed convex hulls is nonexpansive with respect to the Hausdorff metric. However, the proofs presented throw much additional light on the structure of the problem.

4. Approximation by Convex Functions

In this section, we consider the problem of finding a convex function nearest to a given function f. We show that the operator mapping f to the maximal optimal solution of the problem is an optimal Lipschitzian selection operator with $C=2$ and $D=1$.

Let $S=I$ and B be as in Section 3. Let K be the set of all convex functions on I. It is easy to verify that K is a closed convex cone, i.e., K is closed and $\lambda f+\mu h\in K$ whenever $f, h\in K$ and $\lambda\geqslant 0$, $\mu\geqslant 0$. We let $K_f=K$ in (1.2) and rewrite (1.2) as

$$\Delta(f)=\|f-g\|=\inf\{\|f-k\|: k\in K\}. \tag{4.1}$$

The following theorem is essentially a restatement of parts of Theorems 2.1 and 3.1 of [19].

THEOREM 4.1. *Let $f\in B$. There exists an optimal solution $f'\in K$ to the problem* (4.1) *with the property that $f'\geqslant g$ for all optimal solutions g to* (4.1). *Let $\bar{f}$ be the greatest convex minorant of f and $\bar{\Delta}(f)=\|f-\bar{f}\|$. Then the following holds*:

$$\Delta(f)=(\tfrac{1}{2})\,\bar{\Delta}(f), \tag{4.2}$$

$$f'(s)=\bar{f}(s)+\Delta(f)=\bar{f}(s)+(\tfrac{1}{2})\,\bar{\Delta}(f) \qquad \textit{for all } s\in I, \tag{4.3}$$

where obviously $\Delta(f)=\|f-f'\|$.

Such a solution f' is called the maximal optimal solution to the problem. The above theorem shows that it is the greatest convex minorant shifted upward through a certain distance. For the function f_0 defined by (3.3) it is easy to see that $f_0'(s)=2s$, $0\leqslant s\leqslant 1$, and when $f=f_0$,

$$G_f=\{g\colon g \text{ is convex and } 0\leqslant g(s)\leqslant f_0'(s) \text{ for all } 0\leqslant s\leqslant 1\}.$$

We now state our main result for the operator $f\to f'$.

THEOREM 4.2. *Define $T\colon B\to B$ by $T(f)=f'$ where f' is the maximal optimal solution to* (4.1). *Then*

$$\|T(f)-T(h)\|\leqslant\|f-h\|+|\Delta(f)-\Delta(h)| \qquad \textit{for all } f, h\in B. \tag{4.4}$$

Consequently,

$$\|T(f)-T(h)\|\leqslant\|f-h\| \qquad \textit{if} \quad \Delta(f)=\Delta(h). \tag{4.5}$$

Furthermore,

$$|\Delta(f)-\Delta(h)|\leqslant\|f-h\| \qquad \textit{for all } f, h\in B, \tag{4.6}$$

and

$$\|T(f)-T(h)\| \leqslant 2\,\|f-h\| \qquad \textit{for all } f, h \in B. \tag{4.7}$$

T is an optimal Lipschitzian selection operator with $C=2$ and $D=1$ in (1.3) *and* (1.5), *respectively.*

Proof. This follows from Theorems 4.1 and 3.1. Clearly, (4.4) is a consequence of (4.3) and (3.4). Equality (4.2) combined with (3.5) gives (4.6). Again, (4.4) and (4.6) give (4.7).

We now establish the optimality of T. We show that $C=2$ is optimal in the sense of (1.4). Define a sequence of functions f_n, $n=1, 2, \ldots$, on $I=[0, 1]$ by

$$\begin{aligned} f_n(s) &= -1+2ns, & 0 \leqslant s < 1/n,\\ &= 1, & 1/n \leqslant s \leqslant 1. \end{aligned}$$

It is easy to verify that $f_n'(s)=2s-1/n$, $0 \leqslant s \leqslant 1$, and $G_f=\{f_n'\}$ when $f=f_n$. Consequently, for any selection operator T' we must have $T'(f_n)=f_n'$. Let $h=0$ on I, then $h'=0$ on I, $G_h=\{h'\}$, and $T'(h)=h'$. Clearly $\|f_n-h\|=1$, $\|f_n'-h'\|=2-1/n$, $\Delta(f_n)=1-1/n$, and $\Delta(h)=0$. Hence (1.3) with $T=T'$ shows that $C \geqslant 2$. But since (4.7) shows that $C \leqslant 2$, the optimality of $C=2$ and hence of T is established. Inequality (4.6) shows that $D \leqslant 1$. The optimality of $D=1$ in the sense of (1.6) follows immediately from Theorem 3.1 and the fact that $\Delta(f)=\bar{\Delta}(f)/2$. The proof of the theorem is now complete.

We now make two remarks. Observing (4.5), we define

$$C'=\inf\sup\{\|T'(f)-T'(h)\|/\|f-h\| : f, h \in B, f \neq h, \Delta(f)=\Delta(h)\},$$

where the infimum is taken over all T'. We assert that $C'=1$ and the infimum is attained when $T'=T$. Thus T is also optimal in this restricted sense. To prove this assertion, let $f(s)=1$ on I and h be identically zero as before. Then $f'=f$ and $h'=h$. Consequently, $\Delta(f)=\Delta(h)=0$ and $\|f-h\|=1$. For any selection operator T', we must have $T'(f)=f$ and $T'(h)=h$. It follows that $C' \geqslant 1$. Now (4.5) shows that $C'=1$ and hence the assertion is established.

Our second remark pertains to the derivation of (4.6). From first principles, we obtained (3.5), and then using (4.2) established (4.6). However, results of type (4.6) are more general. Indeed, if F is a nonempty subset of a normed linear space X, then for all x, y in X we have

$$|E(x, F)-E(y, F)| \leqslant \|x-y\|, \tag{4.8}$$

where

$$E(x, F)=\inf\{\|x-z\| : z \in F\}.$$

See, for example, [12, p. 17]. Hence, (4.6) follows at once from (4.8) with $X = B$ and $F = K$. Then (4.6) and (4.2) establish (3.5).

5. Generalized Isotone Optimization

In this section, we determine an optimal Lipschitzian selection operator for the problem of generalized isotone optimization and also construct a class of nonoptimal Lipschitzian selection operators with different values of C.

Let S be any set with partial order $\leqslant$. A partial order is a relation $\leqslant$ on S satisfying (i) reflexivity, i.e., $s \leqslant s$ for all s in S, and (ii) transitivity, i.e., if s, t, u are in S and $s \leqslant t$, $t \leqslant u$ then $s \leqslant u$ [6, p. 4]. A partial order is said to be antisymmetric if s, t are in S, and $s \leqslant t$, $t \leqslant s$ then $s = t$. For sake of generality we do not include the condition of antisymmetry in the definition of the partial order as is often done. Let B be the set of all bounded functions on S with the uniform norm (1.1). Let P and Q be two arbitrary subsets of S. Given f in B, the generalized isotone optimization problem is to minimize $\|f - k\|$ over all k in B satisfying

$$k(s) \leqslant k(t) \qquad \text{for all } s \leqslant t, \tag{5.1}$$

$$k(s) \geqslant 0 \qquad \text{for all } s \in P, \tag{5.2}$$

$$k(s) \leqslant 0 \qquad \text{for all } s \in Q. \tag{5.3}$$

We let K denote all functions k in B satisfying (5.1)–(5.3). It is easy to verify that K is a closed convex cone. It is never empty since the function which is identically zero on S is in K. As before, we denote by g an optimal solution to the above problem so that

$$\Delta(f) = \|f - g\| = \inf\{\|f - k\| : k \in K\}.$$

We observe that the above problem allows for equality constraints. Since the partial order is not necessarily antisymmetric, we may let $s \leqslant t$ and $t \leqslant s$, where $s \neq t$. In such a case, (5.1) shows that $k(s) = k(t)$. In addition, if $P \cap Q$ is not empty then $k(s) = 0$ for all s in $P \cap Q$. Some applications of this problem are pointed out in [20] where a weighted uniform norm is considered. We remark here that if $\lambda(s)$, $s \in S$ are real numbers with $|\lambda(s)| = 1$ for all s, then the above problem is equivalent to the one obtained by replacing (5.1) by

$$\lambda(s)\, k(s) \leqslant \lambda(t)\, k(t) \qquad \text{for all } s \leqslant t. \tag{5.4}$$

To see this let $f_1(s) = \lambda(s)\, f(s)$ for all s in S and all f in B. Define

$$P_1 = \{P \cap \{s \in S : \lambda(s) > 0\}\} \cup \{Q \cap \{s \in S : \lambda(s) < 0\}\},$$

$$Q_1 = \{P \cap \{s \in S : \lambda(s) < 0\}\} \cup \{Q \cap \{s \in S : \lambda(s) > 0\}\}.$$

Then our new problem with constraints (5.4), (5.2), and (5.3) is equivalent to minimizing $\|f_1 - k_1\|$ subject to (5.1)–(5.3), where k, P, and Q are respectively replaced by k_1, P_1, and Q_1. This observation immediately shows that approximation by odd functions is a special case of our problem. A real function k defined on an interval $I = [-a, a]$, where $a > 0$, is odd if $k(-s) = -k(s)$ for all s in $[0, a]$. Indeed, in such a case define the partial order $\leqslant$ on S by the following: $s \leqslant s$ for all $s \in S$, $-s \leqslant s$, and $s \leqslant -s$ for all $s \in (0, a]$. Again, we let $P = Q = \{0\}$, $\lambda(s) = 1$ for all s in $[-a, 0]$ and $\lambda(s) = -1$ for all s in $(0, a]$.

Before deriving the Lipschitz condition for our problem, we introduce some notation and state relevant results from [20]. For any subset E of S, we define its indicator function χ_E by

$$\chi_E(s) = 1, \quad \text{if} \quad s \in E,$$
$$= 0, \quad \text{otherwise.}$$

We adopt the convention $0 \cdot \infty = 0$. It is convenient to view the partial order $\leqslant$ as a set of ordered pairs [16, p. 22], viz.

$$\Gamma = \{(s, t) \in S \times S \colon s, t \in S, s \leqslant t\}.$$

Let

$$P_0 = \{t \in S \colon s \leqslant t \text{ for some } s \in P\},$$

$$Q_0 = \{s \in S \colon s \leqslant t \text{ for some } t \in Q\}.$$

A subset E of S is called an upper (lower) set if an element s in S is also in E whenever there exists a t in E with $t \leqslant s$ ($s \leqslant t$). Clearly, P_0 (Q_0) is the smallest upper (lower) set containing $P(Q)$. Given f in B, let

$$\theta(f) = \max\{(1/2) \sup_{(s,t) \in \Gamma} \{f(s) - f(t)\}, \sup_{s \in P_0}\{-f(s)\}, \sup_{s \in Q_0}\{f(s)\}\}. \tag{5.5}$$

Define real valued functions f_* and f^* on S by

$$f_*(s) = \sup_{\{t: t \leqslant s\}} (\max\{f(t) - \theta(f), -\infty(1 - \chi_P(t))\}), \qquad s \in S, \tag{5.6}$$

$$f^*(s) = \inf_{\{u: s \leqslant u\}} (\min\{f(u) + \theta(f), \infty(1 - \chi_Q(t))\}), \qquad s \in S. \tag{5.7}$$

The following theorem is essentially a restatement of Theorems 2.1 and 2.2 of [20].

THEOREM 5.1. *For the generalized isotone optimization problem the following holds*:

(i) *Minimum distance duality*,

$$\theta(f) = \Delta(f) = \min\{\|f - k\| : k \in K\} \qquad \textit{for all } f \in B. \tag{5.8}$$

(ii) *Optimal solutions: Both* f_* *and* f^* *are elements of* K *and are optimal solutions to the problem, i.e.*,

$$\Delta(f) = \|f - f_*\| = \|f - f^*\|.$$

Furthermore, $f_* \leqslant f^*$ *and any* k *in* K *is an optimal solution to the problem if and only if* $f_* \leqslant k \leqslant f^*$.

We now state our main result of this section. For $f \in B$ define $f' = (f_* + f^*))/2$. Since $f_* \leqslant f' \leqslant f^*$ above theorem shows that f' is an optimal solution to our problem and $\Delta(f) = \theta(f) = \|f - f'\|$.

THEOREM 5.2. *Define* $T: B \to B$ *by* $T(f) = f'$. *Then*

$$\|T(f) - T(h)\| \leqslant \|f - h\| \qquad \textit{for all } f, h \in B, \tag{5.9}$$

and

$$|\Delta(f) - \Delta(h)| \leqslant \|f - h\| \qquad \textit{for all } f, h \in B. \tag{5.10}$$

If K *is not identically equal to the zero function on* S, *then* T *is an optimal Lipschitzian selection operator with* $C = 1$ *in* (1.3).

Proof. Since $\Delta(f) = \theta(f)$ and $\Delta(h) = \theta(h)$, using (5.5) one may easily verify that (5.10) holds. Alternatively, (5.10) is an immediate consequence of (4.8). We now show (5.9). Let $s \in S$ and $\varepsilon > 0$. We show that $|f'(s) - h'(s)| \leqslant \|f - h\|$. By the definition of f_* and h^*, there exists $t \leqslant s$ and $u \geqslant s$ such that

$$f_*(s) \leqslant \max\{f(t) - \theta(f), -\infty(1 - \chi_P(t))\} + \varepsilon \tag{5.11}$$

$$h^*(s) \geqslant \min\{h(u) + \theta(h), \infty(1 - \chi_Q(t))\} - \varepsilon. \tag{5.12}$$

Again by the definitions of f^* and h_* we have

$$f^*(s) \leqslant \min\{f(u) + \theta(f), \infty(1 - \chi_Q(u))\}, \tag{5.13}$$

$$h_*(s) \geqslant \max\{h(t) - \theta(h), -\infty(1 - \chi_P(t))\}. \tag{5.14}$$

For convenience, we let

$$m(f) = \max\{f(t) - \theta(f), -\infty(1 - \chi_P(t))\}$$
$$+ \min\{f(u) + \theta(f), \infty(1 - \chi_Q(u))\}$$

and

$$m(h) = \max\{h(t) - \theta(h), -\infty(1 - \chi_P(t))\}$$
$$+ \min\{h(u) + \theta(h), \infty(1 - \chi_Q(u))\}.$$

Clearly, (5.11) and (5.13) show that $f'(s) \leqslant (\frac{1}{2})\, m(f) + \varepsilon/2$. Similarly, (5.12) and (5.14) give $h'(s) \geqslant (\frac{1}{2})\, m(h) - \varepsilon/2$. Consequently,

$$f'(s) - h'(s) \leqslant (1/2)(m(f) - m(h)) + \varepsilon. \tag{5.15}$$

We show that $f'(s) - h'(s) \leqslant \| f - h\| + \varepsilon$. We consider four cases:

(i) Suppose that $t \notin P$ and $u \notin Q$. In this case $\chi_P(t) = \chi_Q(u) = 0$ and consequently $m(f) = f(t) + f(u)$, $m(h) = h(t) + h(u)$. Hence (5.15) gives

$$f'(s) - h'(s) \leqslant (\tfrac{1}{2})(f(t) - h(t) + f(u) - h(u)) + \varepsilon \leqslant \| f - h\| + \varepsilon.$$

(ii) Suppose that $t \in P$ and $u \in Q$. Then since $t \leqslant s \leqslant u$ we have that $t \in Q_0$ and $u \in P_0$. By the definition of $\theta(f)$, we have $\theta(f) \geqslant f(t)$, $\theta(f) \geqslant -f(u)$. Similar conclusions hold for $\theta(h)$. Again, $\chi_P(t) = \chi_Q(u) = 1$. Hence $m(f) = m(h) = 0$. From (5.15) we see that $f'(s) - h'(s) \leqslant \varepsilon$.

(iii) Suppose that $t \in P$, $u \notin Q$. Then

$$m(f) = \max\{f(t) - \theta(f), 0\} + f(u) + \theta(f)$$
$$= \max\{f(t) + f(u), f(u) + \theta(f)\},$$

Similarly,

$$m(h) = \max\{h(t) + h(u), h(u) + \theta(h)\}.$$

Now we use the following inequality,

$$\max\{a_1, a_2\} - \max\{b_1, b_2\} \leqslant \max\{a_1 - b_1, a_2 - b_2\}$$

to conclude that

$$m(f) - m(h) \leqslant \max\{f(t) - h(t) + f(u) - h(u), f(u) - h(u) + \theta(f) - \theta(h)\}.$$

Since $\theta(f) = \Delta(f)$, using (5.10) we find that $m(f) - m(h) \leqslant 2\, \| f - h\|$. Now (5.15) shows that $f'(s) - h'(s) \leqslant \| f - h\| + \varepsilon$.

(iv) Suppose that $t \notin P$ and $u \in Q$. The proof of this case is similar to that of case (iii).

We have thus shown in all cases that $f'(s) - h'(s) \leqslant \| f - h\|$. A symmetric argument completes the proof of $|f'(s) - h'(s)| \leqslant \| f - h\|$. We have thus established (5.9).

We now show that T is optimal. If K is not identically equal to the zero function, then since K is a cone, it must have two distinct elements, say f and h. For any selection operator T' we must have $T'(f)=f$ and $T'(h)=h$. Hence (1.3) with $T=T'$ shows that $C\geqslant 1$. But since (5.9) shows $C\leqslant 1$, the optimality of $C=1$ in the sense of (1.4) and hence of T is established. The proof of the theorem is now complete.

We make one remark. Clearly, (5.10) shows that $D\leqslant 1$ in (1.5). However, in the generality of the statement of Theorem 5.2, one cannot conclude that $D=1$. Indeed, if $P=Q=\varnothing$ and $s\leqslant t$ if $s=t$, then we have $K=B$. Then $\Delta(f)=0$ for all f in B. Thus $D=0$.

We now construct a class of nonoptimal Lipschitzian selection operators T_λ with different values of C. Indeed, for $0\leqslant\lambda\leqslant 1$ let $T_\lambda(f)=\lambda f_*+(1-\lambda)f^*$. Since $f_*\leqslant T_\lambda(f)\leqslant f^*$, by Theorem 5.1, $T_\lambda(f)$ is an optimal solution of the problem for all $0\leqslant\lambda\leqslant 1$. The following result may be established by arguments similar to those used in the proof of Theorem 5.2.

$$\|T_\lambda(f)-T_\lambda(h)\|\leqslant(1+|1-2\lambda|)\,\|f-h\| \qquad \text{for all } f,\, h\in B. \tag{5.16}$$

When $\lambda=\frac{1}{2}$, T_λ equals the operator T in Theorem 5.2. The following example will show that the constant $(1+|1-2\lambda|)$ in (5.16) cannot be reduced. Let $S=I=[0,1]$ with usual order on reals, $P=Q=\varnothing$. Define

$$\begin{aligned} f(s) &= -1, && s=0,\tfrac{1}{2},\\ &= 1 && \text{otherwise.}\end{aligned}$$

Then

$$\begin{aligned} f_*(s) &= -2, && s=0,\\ &= 0 && \text{otherwise,}\end{aligned}$$

and

$$\begin{aligned} f^*(s) &= 0, && 0\leqslant s\leqslant\tfrac{1}{2},\\ &= 2 && \text{otherwise.}\end{aligned}$$

Let h be the identically zero function, then equality holds in (5.16) for all values of $0\leqslant\lambda\leqslant 1$.

ACKNOWLEDGMENT

The author is indebted to the referee for constructive suggestions.

REFERENCES

1. J. L. BENTLEY AND M. I. SHAMOS, Divide and conquer for linear expected time, *Inform. Process. Lett.* **2** (1978), 87–91.
2. A. BYKAT, Convex hull of a finite set points in two dimensions, *Inform. Process. Lett.* **7** (1978), 296–298.
3. B. L. CHALMERS AND G. D. TAYLOR, Uniform approximation with constraints, *Jahresber. Deutsch. Math.-Verein* **81** (1979), 49–86.
4. R. B. DARST AND R. HUOTARI, Best L_1-approximation of bounded, approximately continuous functions on $[0, 1]$ by nondecreasing functions, *J. Approx. Theory* **43** (1985), 178–189.
5. R. B. DARST AND S. SAHAB, Approximation of continuous and quasi-continuous functions by monotone functions, *J. Approx. Theory* **38** (1983), 9–27.
6. N. DUNFORD AND J. T. SCHWARTZ, "Linear Operators, Part I," Interscience, New York, 1958.
7. A. A. GOLDSTEIN, "Constructive Real Analysis," Harper & Row, New York, 1967.
8. R. L. GRAHAM, An efficient algorithm for determining the convex hull of a finite planar set, *Inform. Process. Lett.* **1** (1972), 132–133.
9. R. L. GRAHAM AND F. F. YAO, Finding the convex hull of a simple polygon, *J. Algorithms* **4** (1983), 324–331.
10. C. HILDRETH, Point estimates of ordinates of concave functions, *J. Amer. Statist. Assoc.* **49** (1954), 598–619.
11. M. D. INTRILIGATOR, "Mathematical Optimization and Economic Theory," Prentice–Hall, Englewood Cliffs, N.J., 1971.
12. N. P. KORNEICHUK, "Extremal Problems in Approximation Theory," Nauka, Moscow, 1976.
13. D. LANDERS AND L. ROGGE, Isotonic approximation in L_s, *J. Approx. Theory* **31** (1981), 199–223.
14. F. P. PREPARATA AND S. J. HONG, Convex hulls of finite set of points in two and three dimensions, *Comm. ACM* **20** (1977), 87–93.
15. A. W. ROBERTS AND D. E. VARBERG, "Convex Functions," Academic Press, New York, 1973.
16. H. L. ROYDEN, "Real Analysis," 2nd ed., Macmillan Co., New York, 1968.
17. R. SMARZEWSKI, Determination of Chebyshev approximations by nonlinear admissible subsets, *J. Approx. Theory* **37** (1983), 69–88.
18. V. A. UBHAYA, Isotone optimization, I, II, *J. Approx. Theory* **12** (1974), 146–159, 315–331.
19. V. A. UBHAYA, An $O(n)$ algorithm for discrete n-point convex approximation with applications to continuous case, *J. Math. Anal. Appl.* **72** (1979), 338–354.
20. V. A. UBHAYA, Generalized isotone optimization with applications to starshaped functions, *J. Optim. Theory Appl.* **29** (1979), 559–571.
21. V. A. UBHAYA, Linear time algorithms for convex and monotone approximation, *Comput. Math. Appl.* **9** (1983), 633–643.
22. V. A. UBHAYA, $O(n)$ algorithms for discrete n-point approximation by quasi-convex functions, *Comput. Math. Appl.* **10** (1984), 365–368.
23. V. A. UBHAYA, Quasi-convex optimization, *J. Math. Anal. Appl.*, to appear.