JOURNAL OF APPROXIMATION THEORY 12, 146–159 (1974)

Isotone Optimization, I*

VASANT A. UBHAYA

Department of Computer Science, Washington University, Box 1045
Saint Louis, Missouri 63130

Communicated by John R. Rice

1. INTRODUCTION

Given a function f and a certain cone of functions, the problem of isotone optimization involves determining an element in the cone nearest to f. A suitable norm $\|\cdot\|$ is introduced as a measure of the distance $\|f_1 - f_2\|$ between two functions, f_1 and f_2. The term "isotone" originates from the fact that the cone of functions under consideration is indeed the cone of isotone functions on a partially ordered set. Specifically, let X be a partially ordered set with a partial order $\leq$ and let $\mathscr{V} = \mathscr{V}(X)$ be the linear space of all bounded real valued functions defined on X. A function $h \in \mathscr{V}$ is called an isotone function if $h(x) \leqslant h(y)$ whenever $x, y \in X$ and $x \leq y$. Let $\mathscr{M} = \mathscr{M}(X) \in \mathscr{V}$ be the convex cone of isotone functions. Given a $w \in \mathscr{V}$, $w(x) \geqslant \delta > 0$ for all $x \in X$, define a weighted uniform norm $\|\cdot\|_w$ on $\mathscr{V}$ by

$$\|f\|_w = \sup_{x \in X} w(x)\,|f(x)|, \qquad f \in \mathscr{V}. \tag{1.1}$$

The problem under consideration in this article is: Given $f \in \mathscr{V}$, find $g \in \mathscr{M}$, if one exists, such that

$$\|f - g\|_w = \inf_{h \in \mathscr{M}} \|f - h\|_w . \tag{1.2}$$

Note that if $f \notin \mathscr{M}$ then $\inf_{h \in \mathscr{M}} \|f - h\|_w > 0$ and hence arbitrarily close "approximation" to f is not possible. The weight function w is deliberately

* This two-part article is based on a section of the author's dissertation submitted in partial fulfillment of the requirement for the degree of Ph.D at the University of California, Berkeley in 1971. This research was supported by the National Science Foundation and the Office of Naval Research. The material was revised under Grant GK-32712 from the National Science Foundation. For an abstract of a more general version of this problem see *Notices Amer. Math. Soc. 18* (1971), 961.

introduced in (1.1) to take into account the relative importance of values of f on X. If $w(x) = 1$ for all $x \in X$, (1.1) gives the usual uniform or Tchebycheffian norm well known in approximation theory. Of course, norms other than $\|\cdot\|_w$ may be considered in (1.2). The solution of the problem then will evidently depend upon the norm under consideration.

In Section 2 we first consider the case in which $X = [a, b]$, a closed interval of the real line, which is a totally ordered set. In this case $\mathscr{M}$ becomes the class of monotone (nondecreasing) functions on $[a, b]$. In remarks in this section we indicate how to extend the results of this case to an arbitrary partially ordered set. In Theorem 1, we establish a duality result which gives the value of $\inf_{h \in \mathscr{M}} \| f - h \|_w$ in terms of f and w only. We also show the existence of a solution g satisfying (1.2) and give explicit expressions for the set of all such solutions, which is easily seen to be a convex set. In Theorem 2 we establish properties of this solution set. Our main result is Theorem 3, which states that, if f is continuous on $[a, b]$ and not itself nondecreasing, we can find an infinitely differentiable function g in $\mathscr{M}$ satisfying (1.2).

In [8] we considered a more abstract version of this problem on $L_\infty(X, \Sigma, \mu)$ where X is a totally ordered set, (X, Σ, μ), a complete positive measure space and L_∞, the space of μ-essentially bounded μ-measurable real functions defined on X. The setting of the case of the bounded functions considered in this article has, owing to its more restrictive nature, a richer structure than the general abstract version, and here as well as in [9] we pursue the investigations further to analyze this structure. With reference to the duality result contained in Theorem 1 we remark here that the duality principle in linear spaces points out a correspondence between an extremum problem on the space and an extremum problem on the dual space. For a detailed treatment of duality relationships encountered when approximating elements from those in a given convex set see Ubhaya [10] and other references therein.

A class of problems similar to the one considered in this article has appeared frequently in the literature. Here a function defined on $\mathscr{V}$ and satisfying certain conditions is minimized on $\mathscr{M}$ instead of the norm $\|\cdot\|_w$. See [5, 9, 13, 14] and other references therein. A particular example of this problem occurs when the L_p norm, $1 \leqslant p < \infty$, is used in (1.2) instead of the norm $\|\cdot\|_w$. These problems are motivated mainly because of their applications to statistical analysis involving restricted maximum likelihood estimation. In [9] we shall elaborate more on these problems and show that the problem with the L_p norm has a definite relationship to our problem defined by (1.2). Still another class of problems involving approximation from finite dimensional spaces called the problem of monotone polynomial approximation introduced by Shisha [6] and investigated further by G. G. Lorentz and Zeller [2], R.A. Lorentz [3], Ubhaya [12] and others, has aroused interest in the literature. All these problems involve a common conceptapproximation from a convex cone of functions—predominantly the convex cone of isotone or nondecreasing functions.

2. Main Results

We consider the case when $X = [a, b]$, a closed interval of the real line. Let $\mathscr{V}$, $\mathscr{M}$ and $\|\cdot\|_w$ be as defined in Section 1 with the modification that $X = [a, b]$. In addition let $\mathscr{C} = \mathscr{C}[a, b] \subset \mathscr{V}$ be the linear space of continuous functions and $\mathscr{C}^\infty = \mathscr{C}^\infty[a, b] \subset \mathscr{C}$, the linear space of infinitely differentiable functions. For $\mathscr{U} \subset \mathscr{V}$ let

$$\rho_{\mathscr{U}}(f) = \inf_{u \in \mathscr{U}} \|f - u\|_w, \qquad f \in \mathscr{V}.$$

The problem of Section 1 takes the form: Given $f \in \mathscr{V}$, find $g \in \mathscr{M}$ such that $\rho_{\mathscr{M}}(f) = \|f - g\|_w$.

Let

$$S = \{(x, y) \in [a, b] \times [a, b]: x, y \in [a, b], x \leqslant y\}.$$

For a fixed f in $\mathscr{V}$ we define the following:

$$\theta = \sup_{(x,y)\in S} \frac{w(x)\,w(y)}{w(x) + w(y)} (f(x) - f(y));$$

$$T = \left\{(x, y): (x, y) \in S, \frac{w(x)\,w(y)}{w(x) + w(y)} (f(x) - f(y)) = \theta\right\};$$

$$P = \bigcup_{(x,y)\in T} [x, y];$$

$$Q = \bigcup_{(x,y)\in T} \{x, y\}; \qquad \text{and}$$

$$m(x, y) = \frac{w(x) f(x) + w(y) f(y)}{w(x) + w(y)}, \qquad x, y \in [a, b].$$

Define also the functions,

$$\underline{g}(x) = \sup_{z\in[a,x]} (f(z) - \theta/w(z)), \qquad x \in [a, b],$$

$$\bar{g}(x) = \inf_{z\in[x,b]} (f(z) + \theta/w(z)), \qquad x \in [a, b].$$

We now state our theorems.

THEOREM 1. *Let $f, w \in \mathscr{V}$. Then we have:*

(A) *The Duality Theorem.*

$$\theta = \sup_{(x,y)\in S} \frac{w(x)\, w(y)}{w(x) + w(y)} (f(x) - f(y)) = \min_{h \in \mathscr{M}} \| f - h \|_w = \rho_{\mathscr{M}}(f). \tag{2.1}$$

Hence $\theta \leqslant \| f \|_w$.

(B) *The Characterization Theorem.*

$$\underline{g}, \bar{g} \in \mathscr{M}, \quad \underline{g} \leqslant \bar{g} \qquad \text{and} \qquad \theta = \rho_{\mathscr{M}}(f) = \| f - \underline{g} \|_w = \| f - \bar{g} \|_w . \tag{2.2}$$

Furthermore, for $g \in \mathscr{M}$

$$\theta = \rho_{\mathscr{M}}(f) = \| f - g \|_w$$

holds if and only if $\underline{g} \leqslant g \leqslant \bar{g}$.

THEOREM 2.

(A) *Let $f, w \in \mathscr{C}$. Then in addition to the results of Theorem* 1, $\underline{g}, \bar{g} \in \mathscr{M} \cap \mathscr{C}$. *Hence*

$$\theta = \rho_{\mathscr{M}}(f) = \rho_{\mathscr{M}\cap\mathscr{C}}(f) = \| f - \underline{g} \|_w = \| f - \bar{g} \|_w . \tag{2.3}$$

If $f \notin \mathscr{M}$ $(\Leftrightarrow \theta > 0)$ *then*

$$P = \bigcup_{k=1}^{s} [c_k , d_k], \tag{2.4}$$

where s is some positive integer,

$$a \leqslant c_k < d_k \leqslant b \qquad \text{for all } k,$$
$$d_k < c_{k+1} , \qquad k = 1, 2,\ldots, s-1$$

and

$$(c_k , d_k) \in T \qquad \text{for all } k.$$

Also, in this case the following holds:

(i) $$\underline{g}(x) = \bar{g}(x) \qquad \text{if and only if} \quad x \in P, \tag{2.5}$$

with

$$\underline{g}(x) = \bar{g}(x) = m(c_k , d_k) \qquad \text{for all} \quad x \in [c_k , d_k], \quad \text{for all } k, \tag{2.6}$$

where

$$m(c_k , d_k) < m(c_{k+1} , d_{k+1}), \qquad k = 1, 2,\ldots, s-1. \tag{2.7}$$

(ii) $$w(x)\, | f(x) - \underline{g}(x)| = w(x)\, | f(x) - \bar{g}(x)| = \theta \quad \text{for all} \quad x \in Q. \tag{2.8}$$

(B) *If f, w are absolutely continuous, then g, $\bar{g}$ are absolutely continuous and other results as in* (A) *hold.*

THEOREM 3. *Let $f, w \in \mathscr{C}, f \notin \mathscr{M}$. Then there exists a $k \in \mathscr{M} \cap \mathscr{C}^\infty$ such that*

$$0 < \theta = \rho_{\mathscr{M}}(f) = \rho_{\mathscr{M} \cap \mathscr{C}^\infty}(f) = \|f - k\|_w .$$

Remarks. (i) The results of Theorem 1 are true for an arbitrary partially ordered set X with order $\leq$ provided we replace S, g and $\bar{g}$ in Theorem 1 by the corresponding quantities defined below:

$$S = \{(x, y) \in X \times X: x, y \in X, x \leq y\},$$

$$g(x) = \sup_{\{z \in X: z \leq x\}} (f(z) - \theta/w(z)), \qquad x \in X,$$

$$\bar{g}(x) = \inf_{\{z \in X: z \geq x\}} (f(z) + \theta/w(z)), \qquad x \in X.$$

The proof for the partially ordered case is similar to that of Theorem 1 given in Section 3. When X is totally ordered, Theorem 1 also follows from the abstract results in Ubhaya [8], however, the version of Theorem 1 for a partially ordered set and Theorems 2 and 3 do not follow from these results. Since a much simpler proof can be given to cover all the cases considered in this article, we give it in Section 3.

A min-max form of a g satisfying (1.2) appears in Ubhaya [11]. It is given by $g(x) = \inf \sup m(y, z) = \sup \inf m(y, z)$, for $x \in X$, where the inf and sup are taken respectively over the sets $\{z \in X : z \geq x\}$ and $\{y \in X : y \leq x\}$. A proof of this result may be given as in [8].

(ii) When $X = \{x_1, x_2, \ldots, x_n\}$ is a finite partially ordered set, the isotone optimization problem has the following linear programming formulation:

Minimize z, subject to

$$w_i(f_i - g_i) \leqslant z, \qquad i = 1, 2, \ldots, n,$$

$$-w_i(f_i - g_i) \leqslant z, \qquad i = 1, 2, \ldots, n,$$

$$g_i \leqslant g_j \quad \text{whenever} \quad x_i, x_j \in X \quad \text{and} \quad x_i \leq x_j .$$

Here, for convenience we have introduced the notation, $f(x_i) = f_i$, $g(x_i) = g_i$ and $w(x_i) = w_i$ for all i. For an introduction to linear programming see Dantzig [1]. Our characterization of the solution of the problem indeed shows how to solve the problem without resorting to some form of matrix inversion necessary in linear programming.

(iii) It is the result of Lemma 1 of Section 3 that $f \in \mathscr{M}$ if and only if $\theta = 0$. Hence from the definitions of g and $\bar{g}$ we conclude that $g = \bar{g} = f$ if and only if $f \in \mathscr{M}$. Therefore if $f \in \mathscr{M} - \mathscr{C}^\infty$, there does not exist $g \in \mathscr{M} \cap \mathscr{C}^\infty$

satisfying $\rho_{\mathscr{M}}(f) = \|f - g\|_w$. Theorem 3, on the other hand, shows that if $f \in \mathscr{C} - \mathscr{M}$ and $w \in \mathscr{C}$, then it is always possible to find such a $g \in \mathscr{M} \cap \mathscr{C}^\infty$.

(iv) To understand the significance of the duality result, Theorem 1, A, in terms of the dual extremum problem, see Ubhaya [10].

3. Proofs

We now proceed to the proofs of theorems stated in Section 2.

3.1. Lemma 1. $\theta = 0$ *if and only if* $f \in \mathscr{M}$.

Proof. If $f \in \mathscr{M}$, then for all $(x, y) \in S$ we have $f(x) \leqslant f(y)$. Since $(x, x) \in S$ for all $x \in [a, b]$, it follows that $\theta = 0$. If $f \notin \mathscr{M}$, then there exists $(x, y) \in S$, $x < y$ such that $f(x) > f(y)$. Hence

$$\theta \geqslant (w(x)\, w(y)/(w(x) + w(y)))(f(x) - f(y)) > 0.$$

3.2. *Proof of Theorem* 1. We first show that $\theta \leqslant \|f - h\|_w$ for all $h \in \mathscr{M}$. Let $(x, y) \in S$ and $\eta = \|f - h\|_w$. Then clearly

$$w(x)\,|f(x) - h(x)| \leqslant \eta,$$
$$w(y)\,|f(y) - h(y)| \leqslant \eta.$$

Now since $h(y) - h(x) \geqslant 0$, we have,

$$\begin{aligned} f(x) - f(y) &\leqslant f(x) - f(y) + h(y) - h(x) \\ &\leqslant |f(x) - h(x)| + |f(y) - h(y)| \\ &\leqslant \eta(1/w(x) + 1/w(y)). \end{aligned}$$

It follows that $\theta \leqslant \|f - h\|_w$.

Clearly $\underline{g}$ and $\bar{g} \in \mathscr{M}$. To prove (A), it now suffices to show that $\theta = \|f - \underline{g}\|_w$. Let $x \in [a, b]$. From the definition of $\underline{g}$ it follows that $\underline{g}(x) \geqslant f(x) - \theta/w(x)$, i.e., $w(x)(\underline{g}(x) - f(x) \geqslant -\theta$. Now let $\epsilon > 0$; then there exists $z \in [a, x]$ such that $\underline{g}(x) \leqslant f(z) - \theta/w(z) + \epsilon$. By the definition of θ we must have,

$$f(z) - \theta/w(z) \leqslant f(x) + \theta/w(x).$$

Hence, $\underline{g}(x) \leqslant f(x) + \theta/w(x) + \epsilon$. It follows that $\underline{g}(x) \leqslant f(x) + \theta/w(x)$, i.e., $w(x)(\underline{g}(x) - f(x)) \leqslant \theta$. Thus $\|f - \underline{g}\|_w = \theta$.

(B) As in the proof for (A), we may show that $\|f - \bar{g}\|_w = \theta$. Suppose now $g \in \mathscr{M}$ and satisfies $\theta = \|f - g\|_w$. Let $x \in [a, b]$. By the definition of $\underline{g}$, given an $\epsilon > 0$, there exists $z \in [a, x]$ such that $\underline{g}(x) \leqslant f(z) - \theta/w(z) + \epsilon$. Now, since $g(z) \geqslant f(z) - \theta/w(z)$ and $g \in \mathscr{M}$, we have

$$\underline{g}(x) \leqslant g(z) + \epsilon \leqslant g(x) + \epsilon.$$

Hence $\underline{g} \leqslant g$ and specializing for $g = \bar{g}$ we have $\underline{g} \leqslant \bar{g}$. We may similarly show that $g \leqslant \bar{g}$. If $\underline{g} \leqslant g \leqslant \bar{g}$, then clearly $\| f - g \|_w = \theta$. The proof of Theorem 1 is now complete.

In what follows we establish a number of lemmas before we prove Theorems 2 and 3. These lemmas will first uncover properties of the sets T, P and Q and subsequently will reveal the structure of $\underline{g}$ and $\bar{g}$.

3.3. *Decomposition of P.*

In this section we assume that $f, w \in \mathscr{C}$ and $\theta > 0$. Note that by Lemma 1, the assumption $\theta > 0$ is equivalent to $f \notin \mathscr{M}$. It follows that T, P and Q are nonempty. We show that there exist a finite number of disjoint closed intervals $[c_k, d_k] \subset [a, b]$, $k = 1, 2, \ldots, s$ such that (2.4) holds.

Clearly $m(x, y)\colon [a, b] \times [a, b] \to R$ is a continuous function. Let

$$\Gamma = \{\gamma\colon \gamma = m(x, y), (x, y) \in T\}.$$

We define an equivalence relation $\sim$ on T by $(x, y) \sim (u, v)$ if and only if $m(x, y) = m(u, v)$, where $(x, y), (u, v) \in T$. Then,

$$T_\gamma = \{(x, y)\colon (x, y) \in T, m(x, y) = \gamma\}, \qquad \gamma \in \Gamma,$$

are equivalence classes.

By continuity of m, Γ is bounded. Define

$$c_\gamma = \inf\{x\colon (x, y) \in T_\gamma\},$$
$$d_\gamma = \sup\{y\colon (x, y) \in T_\gamma\}.$$

Clearly, $c_\gamma < d_\gamma$ since by our assumption $f \notin \mathscr{M}$.

LEMMA 2. *Let* $\gamma, \sigma \in \Gamma$, $(x, y) \in T$. *Then*

(i) $\quad m(c_\gamma, d_\gamma) = \gamma, \qquad (c_\gamma, d_\gamma) \in T_\gamma, \qquad \{c_\gamma, d_\gamma\} \subset Q.$

(ii) $\quad [x, y] \cap [c_\gamma, d_\gamma] \neq \varnothing \Leftrightarrow m(x, y) = \gamma.$

and hence

$$[c_\gamma, d_\gamma] = \bigcup \{[x, y]\colon (x, y) \in T, [x, y] \cap [c_\gamma, d_\gamma] \neq \varnothing\}.$$

(iii) $\quad \gamma \neq \sigma \Leftrightarrow [c_\gamma, d_\gamma] \cap [c_\sigma, d_\sigma] = \varnothing.$

Proof. (i) There exists (x_n, y_n), $(u_n, v_n) \in T$, $n = 1, 2, \ldots$ such that $x_n \to c_\gamma$, $v_n \to d_\gamma$, $m(x_n, y_n) = m(u_n, v_n) = \gamma$, for all n. We then have,

$$f(x_n) - \theta/w(x_n) = f(y_n) + \theta/w(y_n) = \gamma,$$
$$f(u_n) - \theta/w(u_n) = f(v_n) + \theta/w(v_n) = \gamma.$$

Consequently,

$$f(x_n) - \theta/w(x_n) = \gamma = f(v_n) + \theta/w(v_n).$$

Letting $n \to \infty$, by the continuity of f and w we have

$$f(c_\gamma) - \theta/w(c_\gamma) = \gamma = f(d_\gamma) + \theta/w(d_\gamma). \tag{3.1}$$

It follows that

$$(w(c_\gamma)\, w(d_\gamma)/(w(c_\gamma) + w(d_\gamma)))(f(c_\gamma) - f(d_\gamma)) = \theta$$

Hence $(c_\gamma, d_\gamma) \in T$. Now multiplying the first equation in (3.1) by $w(c_\gamma)$ and the second equation by $w(d_\gamma)$, adding and simplifying, we have $m(c_\gamma, d_\gamma) = \gamma$ and hence $\{c_\gamma, d_\gamma\} \subset Q$.

(ii) By the definition of θ, for $a \leqslant x \leqslant y \leqslant b$ we have,

$$f(x) - \theta/w(x) \leqslant f(y) + \theta/w(y). \tag{3.2}$$

If $[c_\gamma, d_\gamma] \cap [x, y] \neq \varnothing$ then we have $x \leqslant d_\gamma$, $c_\gamma \leqslant y$. From (3.1), (3.2) and the definition of θ it follows that

$$\begin{aligned} f(x) - \theta/w(x) &\leqslant f(d_\gamma) + \theta/w(d_\gamma) = \gamma = f(c_\gamma) - \theta/w(c_\gamma) \\ &\leqslant f(y) + \theta/w(y). \end{aligned} \tag{3.3}$$

But since $(x, y) \in T$, (3.2) holds with equality and hence from (3.3)

$$f(x) - \theta/w(x) = f(y) + \theta/w(y) = \gamma,$$

and $m(x, y) = \gamma$. If $(x, y) \in T$ and $m(x, y) = \gamma$ then $c_\gamma \leqslant x < y \leqslant d_\gamma$.

(iii) This follows at once from (i), (ii) and the definition of c_γ, d_γ.

The proof of the lemma is now complete.

Recall our assumption that $f \notin \mathscr{M}$. We now state and prove:

LEMMA 3. (Finiteness of Γ). *Γ is a finite set.*

Proof. Consider (c_γ, d_γ), $\gamma \in \Gamma$. Then

$$f(c_\gamma) - f(d_\gamma) = \theta(1/w(c_\gamma) + 1/w(d_\gamma)) \geqslant 2\theta/(\max_{x \in [a,b]} w(x)) = \theta' > 0.$$

Therefore by uniform continuity of f on $[a, b]$ there exists $\delta > 0$ such that $d_\gamma - c_\gamma > \delta$ for all $\gamma \in \Gamma$. Since the intervals $[c_\gamma, d_\gamma]$ are disjoint by Lemma 2, the result follows.

We write $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_s\}$ where $\gamma_k < \gamma_{k+1}$, $k = 1, \ldots, s-1$, and $[c_{\gamma_k}, d_{\gamma_k}] = [c_k, d_k]$, $k = 1, 2, \ldots, s$. Then (2.4) holds and $m(c_k, d_k) = \gamma_k$.

3.4. *Properties of g and $\bar{g}$*

In this section we establish properties of g and $\bar{g}$.

LEMMA 4. *If $f, w \in \mathscr{C}$ then $g, \bar{g} \in \mathscr{C}$.*

Proof. Using the definition of g we may write for $y > x$,

$$g(y) = \max\{g(x), \max_{z \in [x,y]} (f(z) - \theta/w(z))\}.$$

Hence,

$$g(y) - g(x) = \max\{0, \max_{z \in [x,y]} (f(z) - \theta/w(z) - g(x))\}.$$

But noting that $g(x) \geqslant f(x) - \theta/w(x)$, we conclude that

$$0 \leqslant g(y) - g(x) \leqslant \max\{0, \max_{z \in [x,y]} ((f(z) - \theta/w(z)) - (f(x) - \theta/w(x)))\}.$$

From the continuity of $f - \theta/w$, the continuity of g follows. In a similar manner we may show that $\bar{g} \in \mathscr{C}$.

LEMMA 5. *Assume $f, w \in \mathscr{C}$ and $\theta > 0$, then*

$$g(x) = \bar{g}(x) = \gamma_k \quad \text{for all} \quad x \in [c_k, d_k], \qquad k = 1, 2,..., s,$$
$$w(x)\, |f(x) - \bar{g}(x)| = w(x)\, |f(x) - \bar{g}(x)| = \theta \qquad \text{for all} \quad x \in Q.$$

Proof. We first show that $g(x) = \gamma_k$ for all $x \in [c_k, d_k]$. Letting $\gamma = \gamma_k$, $c_\gamma = c_k$ and $d_\gamma = d_k$ in (3.1) we conclude that

$$f(c_k) - \theta/w(c_k) = \gamma_k = f(d_k) + \theta/w(d_k), \qquad k = 1, 2,..., s.$$

If $a \leqslant z \leqslant d_k$, then by the definition of θ we have,

$$f(z) - \theta/w(z) \leqslant f(d_k) + \theta/w(d_k) = \gamma_k .$$

From the definition of g it follows that $g(x) \leqslant \gamma_k$ for all $x \in [a, d_k]$. But since

$$g(c_k) \geqslant f(c_k) - \theta/w(c_k) = \gamma_k ,$$

and g is nondecreasing, we conclude that $g(x) = \gamma_k$ for all $x \in [c_k, d_k]$. The fact that $\bar{g}(x) = \gamma_k$ for all $x \in [c_k, d_k]$ may be established in a similar manner.

Let $X \in Q$. We show that $w(x)\, |f(x) - g(x)| = \theta$. By the definition of Q, there exists $y \in [a, b]$ such that either $(x, y) \in T$ or $(y, x) \in T$. We consider the case for which $(x, y) \in T$; the other case can be treated similarly. Since Γ is a finite set (Lemma 3), it follows that $m(x, y) = \gamma_k$ for some $k = 1, 2,..., s$. Hence by Lemma 2, $[x, y] \subset [c_k, d_k]$. Since g is nondecreasing we have

$$\gamma_k = g(c_k) \leqslant g(x) \leqslant g(y) \leqslant g(d_k) = \gamma_k .$$

It follows that $g(x) = \gamma_k$ and

$$w(x)|f(x) - g(x)| = w(x)|f(x) - m(x, y)|$$
$$= \frac{w(x)\, w(y)}{w(x) + w(y)}\, |f(x) - f(y)| = \theta,$$

the last equality following from the fact that $(x, y) \in T$. The proof for $\bar{g}$ is similar.

LEMMA 6. *Assume* $f, w \in \mathscr{C}$ *and* $\theta > 0$; *then*

$$\text{(i)} \qquad \bar{g}(x) > \gamma_k \quad \text{if} \quad x > d_k$$

$$\text{(ii)} \qquad g(x) < \gamma_k \quad \text{if} \quad x < c_k .$$

Proof. (i) Suppose that for some $x > d_k$, $\bar{g}(x) = \gamma_k$. Then by the definition of $\bar{g}$ there exists $y \in [x, b]$ such that

$$\gamma_k = f(c_k) - \theta/w(c_k) = f(y) + \theta/w(y).$$

We then have,

$$(w(c_k)\, w(y)/(w(c_k) + w(y)))(f(c_k) - f(y)) = \theta.$$

Hence $(c_k , y) \in T$ and by Lemma 2, $[c_k , y] \subset [c_k , d_k]$ which is a contradiction since $c_k < d_k < y$.

(ii) The proof for this case is similar.

LEMMA 7. *Assume* $f, w \in \mathscr{C}$ *and* $\theta > 0$: *then*

$$g(x) < \bar{g}(x) \qquad \text{for all } x \in [a, c_1) \cup \left(\bigcup_{k=1}^{s-1} (d_k , c_{k+1}) \right) \cup (d_s , b].$$

Proof. If, on the contrary, for some $t \in (d_k , c_{k+1})$, $k = 1, 2,..., s - 1$, $g(t) = \bar{g}(t)$ holds, then by the definitions of g and $\bar{g}$ we must have

$$g(t) = f(u) - \theta/w(u) = f(v) + \theta/w(v) = \bar{g}(t),$$

where $u \in [d_k , t]$, $v \in [t, c_{k+1}]$. Then

$$(w(u)\, w(v)/(w(u) + w(v)))(f(u) - f(v)) = \theta, \qquad u < v.$$

Hence, $(u, v) \in T$ and by Lemma 2, $[u, v] \subset [c_i , d_i]$ for some i, which is a contradiction.

The same procedure is applicable to the intervals $[a, c_1)$ and $(d_s , b]$.

LEMMA 8. *Assume* $f, w \in \mathscr{C}$ *and* $\theta > 0$. *Define* $d_0 = a - 1$ *and* $c_{s+1} = b + 1$.

(i) *There exists* d_k' *satisfying* $d_k < d_k' < c_{k+1}$, $k = 1, 2,..., s$ *such that* $g(x) = g(d_k) = \gamma_k$ *for all* $x \in [d_k , d_k'] \cap [a, b]$ *and* $g(x) > \gamma_k$ *for all* $x \in (d_k', b + 1] \cap [a, b]$.

(ii) *There exists* c_k' *satisfying* $d_{k-1} < c_k' < c_k$, $k = 1, 2,..., s$ *such that* $\bar{g}(x) = \bar{g}(c_k) = \gamma_k$ *for all* $x \in [c_k', c_k] \cap [a, b]$ *and* $\bar{g}(x) < \gamma_k$ *for all* $x \in [a - 1, c_k') \cap [a, b]$.

Proof. Note that

$$\begin{aligned} \gamma_k &= \underline{g}(c_k) = f(c_k) - \theta/w(c_k) < f(c_k) + \theta/w(c_k), \qquad k = 1, 2,..., s, \\ \gamma_k &= \bar{g}(d_k) = f(d_k) + \theta/w(d_k) > f(d_k) - \theta/w(d_k), \qquad k = 1, 2,..., s. \end{aligned} \tag{3.4}$$

Define

$$E_k = \{x \in [d_k, c_{k+1}] \cap [a, b]: f(x) - \theta/w(x) > \gamma_k\}, \qquad k = 1, 2,..., s,$$

$$F_k = \{x \in [d_{k-1}, c_k] \cap [a, b]: f(x) + \theta/w(x) < \gamma_k\}, \qquad k = 1, 2,..., s,$$

$$d_k' = \begin{cases} \inf E_k, & \text{if } E_k \neq \varnothing, \\ b + 1/2, & \text{otherwise,} \end{cases}$$

$$c_k' = \begin{cases} \sup F_k, & \text{if } F_k \neq \varnothing, \\ a - 1/2, & \text{otherwise.} \end{cases}$$

Note that $E_k \neq \varnothing$, $k = 1, 2,..., s - 1$ and $F_k \neq \varnothing$, $k = 2, 3,..., s$. By the continuity of f and w, (3.4) and the fact that $\gamma_k < \gamma_{k+1}$ we have,

$$\gamma_k = f(d_k') - \theta/w(d_k'), \qquad \text{if } E_k \neq \varnothing,$$

$$\gamma_k = f(c_k') + \theta/w(c_k'), \qquad \text{if } F_k \neq \varnothing.$$

The assertions of the lemma then follow from the continuity of f, w and the definitions of $\underline{g}$ and $\bar{g}$.

LEMMA 9. *If f, w are absolutely continuous, then $\underline{g}$ and $\bar{g}$ are absolutely continuous.*

Proof. We show that $\underline{g}$ is absolutely continuous; the proof for $\bar{g}$ is similar. The hypothesis of the theorem implies that $h = f - \theta/w$ is absolutely continuous and given $\epsilon > 0$ there exists $\delta > 0$ such that for every finite collection $\{(x_i, y_i)\}$ of nonoverlapping intervals $(x_i, y_i) \subset [a, b]$ with $\Sigma_i \mid y_i - x_i \mid < \delta$ we have

$$\Sigma_i \mid h(y_i) - h(x_i) \mid < \epsilon.$$

Suppose now that $\{(x_i, y_i)\}$ is a finite collection of nonoverlapping intervals such that $(x_i, y_i) \subset [a, b]$ for all i. We show that there exists a finite collection $\{(x_i', y_i')\}$ of nonoverlapping intervals such that $(x_i', y_i') \subset (x_i, y_i)$ for all i and

$$\mid \underline{g}(x_i) - \underline{g}(y_i) \mid \leqslant \mid h(x_i') - h(y_i') \mid. \tag{3.5}$$

Absolute continuity of $\underline{g}$ will then follow. Suppose $\underline{g}(x_i) = \underline{g}(y_i)$; then let

$x_i' = x_i$ and $y_i' = y_i$. Clearly (3.5) holds. Now suppose that $g(x_i) < g(y_i)$. Since $g(x) = \max_{z\in[a,x]} h(z)$ we have

$$g(y_i) = \max\{g(x_i), \max_{z\in[x_i,y_i]} h(z)\}.$$

Therefore, there exists y_i' satisfying $x_i < y_i' \leqslant y_i$ such that $g(y_i) = h(y_i')$. Again, $h(z) \leqslant g(x_i) < g(y_i)$ for all $z \in [a, x_i]$. Hence, there exists x_i', $x_i \leqslant x_i' < y_i' < y_i$ such that $g(x_i) = h(x_i')$ and again (3.5) holds in this case.

We now proceed to the proofs of Theorems 2 and 3. They follow directly from the lemmas we have proved.

3.5. *Proof of Theorem* 2. (A) Since $f, w \in \mathscr{C}$, by Lemma 4, g and $\bar{g} \in \mathscr{C}$ and (2.3) holds. Relation (2.4) and other properties of the intervals $[c_k, d_k]$, as well as (2.7) follow from the results of Section 3.3. Relations (2.5), (2.6) and (2.8) are the results of Lemmas 5 and 7.

(B) This part is the result of Lemma 9.

3.6. *Proof of Theorem* 3. Combining the results of Lemma 5 and Lemma 8 we have with $\gamma_0 = g(a)$ and $\gamma_{s+1} = \bar{g}(b)$,

$$g(x) \begin{cases} = \gamma_k & x \in [c_k, d_k'] \cap [a, b] \\ < \gamma_{k+1} \quad \text{and} \quad > \gamma_k & x \in (d_k', c_{k+1}) \cap [a, b]. \end{cases} \tag{3.6}$$

$$\bar{g}(x) \begin{cases} > \gamma_k \quad \text{and} \quad < \gamma_{k+1} & x \in (d_k, c_{k+1}') \cap [a, b], \\ = \gamma_{k+1} & x \in [c_{k+1}', d_{k+1}] \cap [a, b]. \end{cases} \tag{3.7}$$

The expressions (3.6) and (3.7) together with the contents of Lemma 7 allow us to determine a $k \in \mathscr{M} \cap \mathscr{C}^\infty$ such that $k(x) = \gamma_k$ for all $x \in [c_k, d_k]$, $k = 1, 2,..., s$ and $k(x) \in [g(x), \bar{g}(x)]$ for all $x \in [a, b] - \bigcup_{k=1}^{s} [c_k, d_k]$. Thus $g \leqslant k \leqslant \bar{g}$ and it follows from Theorem 1 that $\|f - k\|_w = \theta$.

An explicit expression for such a $k \in \mathscr{M} \cap \mathscr{C}^\infty$ may be obtained by convolving the continuous function g with a Friedrich's Mollifier function (see Morrey [4]). A nonnegative, infinitely differentiable function ϕ defined on the real line is called a mollifier function if its support (i.e., the set of points on which the function is nonzero) is contained in (0, 1) and $\int_0^1 \phi(t)\,dt = 1$. Let $\gamma_0 = g(a)$ and $\gamma_{s+1} = \bar{g}(b)$. Define the sets

$$A_k\{(y - x): \gamma_{k+1} > g(y) = \bar{g}(x) > \gamma_k\}, \qquad k = 0, 1, 2,..., s.$$

Let $\alpha_k = \inf A_k$, with the understanding that $\inf \varnothing = +\infty$. It will be shown later that $\alpha_k > 0$ for all k. Let $\alpha = \min\{1, \{\alpha_k\}_{k=0}^{s}\} > 0$ and define

$$g'(x) = \begin{cases} g(a), & \text{if } x \in (-\infty, a) \\ g(x), & \text{if } x \in [a, b] \\ g(b), & \text{if } x \in (b, \infty) \end{cases}$$

It is easy to see that

$$\underline{g}(x) \leqslant \underline{g}'(x + \alpha t) \leqslant \underline{g}'(x + \alpha) \leqslant \underline{g}(x) \tag{3.8}$$

for all $x \in [a, b]$, all $t \in [0, 1]$. We now define

$$k(x) = \int_0^1 \underline{g}'(x + \alpha t)\, \phi(t)\, dt, \qquad x \in [a, b].$$

Then $k \in \mathscr{M}$ and using (3.8) we conclude that $\underline{g}(x) \leqslant k(x) \leqslant \bar{g}(x)$ for all $x \in [a, b]$. Also $k \in \mathscr{C}^\infty$ (see Morrey [4]).

We now show that $\alpha_k > 0$. From (3.6), (3.7) we conclude that $A_k \neq \varnothing$ for $k = 1, 2,..., s - 1$ and we first consider these cases. If $\gamma_{k+1} > \ (y) = \bar{g}(x) > \gamma_k$, then $y \in (d_k', c_{k+1})$, $x \in (d_k, c'_{k+1})$, and $y - x > 0$. If $\alpha_k = 0$, then we can extract convergent sequences $y_n \in (d_k', c_{k+1})$, $x_n \in (d_k, c'_{k+1})$ such that $y_n \to y^* \in [d_k', c_{k+1}]$, $x_n \to x^* \in [d_k, c'_{k+1}]$ and $(y_n - x_n) \to (y^* - x^*) = 0$. Thus $x^* \in [d_k', c'_{k+1}]$. Continuity of $\underline{g}$, $\bar{g}$ gives $\underline{g}(x^*) = \bar{g}(x^*)$, which by (2.5) implies that $x^* \in P$. But since $[d_k', c'_{k+1}] \cap P = \varnothing$, a contradiction results. Hence $\alpha_k > 0$. Now if $A_0 = \varnothing$, then $\alpha_0 = +\infty$. If $A_0 \neq \varnothing$, then there exists y such that $\bar{g}(y) < \gamma_1$. From (3.7) we conclude that $a < c_1'$. If $\alpha_0 = 0$, then as before by considering convergent sequences, a contradiction is reached. Hence $\alpha_0 > 0$. The proof for α_s is similar.

Acknowledgment

The author would like to express his gratitude to Professors Charles B. Morrey, William G. Bade, and Raphael M. Robinson of the University of California, Berkeley, for their encouragement and help which is more fully acknowledged in [7] and also to Richard E. Barlow for pointing out some references of statistical significance on isotonic regression.

References

1. G. B. Dantzig, "Linear Programming and Extensions," Princeton University Press, Princeton, New Jersey, 1963.
2. G. G. Lorentz and K. L. Zeller, Monotone approximation by algebraic polynomials, *Trans. Amer. Math. Soc.* **149** (1970), 1–18.
3. R. A. Lorentz, Uniqueness of best approximation by monotone polynomials, *J. Approximation Theory* **4** (1971), 401–418.
4. C. B. Morrey, Jr., "Multiple Integrals in the Calculus of Variations," Springer-Verlag, Berlin, New York, 1966.
5. T. Robertson and P. Waltman, On estimating monotone parameters, *Ann. Math. Statist.* **39** (1968), 1030–1039.
6. O. Shisha, Monotone approximation, *Pacific J. Math.* **15** (1965), 667–671.

7. V. A. UBHAYA, Isotonic and convex optimization, *Technical Report ORC* 17–13, Operations Research Center, University of California, Berkeley, California, July 1971, or U.S. Department of Commerce, *NTIS Report* 12*B AD*-734 130, Feb. 10, 1972.
8. V. A. UBHAYA, Almost monotone approximation in L_∞, *AMCS Report AM*-72-3, Department of Applied Mathematics and Computer Science, Washington University, St. Louis, Missouri, 1972. *J. Math. Anal. Appl.*, to appear.
9. V. A. UBHAYA, Isotone optimization II, *J. Approximation Theory*, to appear.
10. V. A. UBHAYA, Minimization of seminorms on a topological vector space — duality and characterization theorems, *J. Math. Anal. Appl.* **46** (1974), 635–657. See also Errata and Addenda to this article in *J. Math. Anal. Appl.* **47** (1974), 441–442.
11. V. A. UBHAYA, Approximation by isotone functions—a min-max form of best approximation, *Notices Amer. Math. Soc.* **21** (1974), A–307.
12. V. A. UBHAYA, Moduli of monotonicity and applications to monotone polynomial approximation, *Notices Amer. Math. Soc.* **20** (1973), A–525.
13. C. VAN EEDEN, Testing and estimating ordered parameters of probability distributions, Ph.D. dissertation, University of Amsterdam, Studentendrukkerij Poortpers N.V., Amsterdam, 1958.
14. C. VAN EEDEN, Maximum likelihood estimation of partially or completely ordered parameters, I and II, *Indag. Math.* **19** (1957), 128–136, 201–211.