JOURNAL OF APPROXIMATION THEORY 12, 315–331 (1974)

Isotone Optimization. II*

VASANT A. UBHAYA

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

Communicated by John R. Rice

1. INTRODUCTION

Let X be a partially ordered set with order $\leq$, let $\mathscr{V} = \mathscr{V}(X)$ be the linear space of bounded real functions on X and $\mathscr{M} = \mathscr{M}(X) \subset \mathscr{V}$ the convex cone of isotone functions in $\mathscr{V}$, i.e., functions h satisfying $h(x) \leqq h(y)$ whenever $x, y \in X$, $x \leq y$. Given a weighted uniform norm $\|\cdot\|_w$ on $\mathscr{V}$ defined by

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

where w in $\mathscr{V}$ is a weight function satisfying $w(x) \geqq \delta > 0$ for all $x \in X$, the problem is to find g in $\mathscr{M}$, if one exists, such that

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

We call this problem the problem of isotone optimization with respect to the weighted uniform norm (1.1). Instead of (1.1) we may consider other norms, e.g., the l_p norm, $1 \leqq p < \infty$. Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite partially ordered set. For each p, $1 \leqq p < \infty$, define an l_p norm $\|\cdot\|_w^p$ by

$$\|f\|_w^p = \left(\sum_{i=1}^{n} w_{p,i}\,|f_i|^p\right)^{1/p}, \tag{1.3}$$

where $f = \{f_i\}_{i=1}^n$ is a function on X and $w_p = \{w_{p,i}\}_{i=1}^n > 0$ is a given weight function. Since X is finite, we identify any function f with the sequence $\{f_1, f_2, \ldots, f_n\}$ where $f_i = f(x_i)$ and for convenience write $f = \{f_i\}_{i=1}^n$. The class $\mathscr{M}$, of isotone functions in this case is the set of all $h = \{h_i\}_{i=1}^n$ on X satisfying

$$h_i \leqq h_j \text{ whenever } x_i, x_j \in X \text{ and } x_i \leq x_j . \tag{1.4}$$

* Part I appears in this Journal **12**, 146–159 (1974). Both parts are 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.

The problem of isotone optimization with respect to the l_p norm is: given $\{f_i\}_{i=1}^n$ find $g_p = \{g_{p,i}\}_{i=1}^n$ in $\mathscr{M}$, if one exists, such that

$$\| f - g_p \|_w^p = \inf_{h \in \mathscr{M}} \| f - h \|_w^p . \tag{1.5}$$

In [11] we considered the problem (1.2) and characterized the set of all its solutions. In this article we investigate the problem (1.2) further and show its relationship to the problem (1.5). In Section 2 we show that when X is finite, under certain conditions on the weight functions w_p, the solution of (1.5) converges as $p \to \infty$ to a solution of problem (1.2) for some $w = \{w_i\}_{i=1}^n$. In Section 3 we point out a norm reducing property of a particular solution of (1.2) when $w(x) = 1$ for all $x \in X$. Specifically if $f_i \in \mathscr{V}$, $i = 1, 2$ and f_i^0, $i = 1, 2$ be the corresponding particular solutions of (1.2), we show that $\| f_1^0 - f_2^0 \|_w \leqq \| f_1 - f_2 \|_w$ holds. In Section 4 we investigate the differentiability properties of the solutions of (1.2) when $X = [a, b]$, a closed interval of the real line and in Section 5 we construct algorithms to compute these solutions and establish relevant rates of convergence.

The problem (1.5) arises in certain aspects of statistical analysis involving the restricted maximum likelihood estimation. To give a simple example, maximizing the joint density of n independent normal distributions $N(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots, n$ with an ordering restriction on the means μ_i is the same as solving an isotone optimization problem with respect to the l_2 norm. See [2, 10]. Owing to the applications to statistics this problem and a more general version which involves minimization of a function defined on $\mathscr{V}$ and satisfying certain conditions, are extensively investigated. For a history of the problem see [2]. See also [3], [7], and [12]. The solution $g_p = \{g_{p,i}\}_{i=1}^n$ of the problem (1.5) for $1 < p < \infty$, is known to be given by (See [12]).

$$\begin{aligned} g_{p,i} &= \max_{\{U: i \in U\}} \min_{\{L: i \in L\}} u_p(L \cap U) \\ &= \min_{\{L: i \in L\}} \max_{\{U: i \in U\}} u_p(L \cap U), \qquad \text{for all } i, \end{aligned} \tag{1.6}$$

where L and U are lower and upper sets respectively and unique $u_p(L \cap U)$ satisfies

$$\sum_{i \in L \cap U} w_{p,i} \, | f_i - u_p(L \cap U) |^p \leqq \sum_{i \in L \cap U} w_{p,i} \, | f_i - u |^p \tag{1.7}$$

for all real u. (We call $L \subset X$ a lower set if $x_i \in L$ and $x_j \in X$, $x_j \leq x_i$ implies that $x_j \in L$. Similarly $U \subset X$ is an upper set if $x_i \in U$ and $x_j \in X$, $x_j \geq x_i$ implies that $x_j \in U$.) When $p = 2$ it is easy to see that (1.7) gives

$$u_2(L \cap U) = \Big(\sum_{i \in L \cap U} w_{2,i} f_i \Big) \Big/ \Big(\sum_{i \in L \cap U} w_{2,i} \Big)$$

and from (1.6) it may be seen that g_2 has an elegant expression. We shall use these expressions in Section 2.

2. Convergence Properties of the Isotone Optimization Problem with Respect to the l_p Norm

Let $X = \{x_1, x_2, ..., x_n\}$ be finite partially ordered. In this case the norm (1.1) takes the form

$$\|f\|_w = \max_{1 \leqq i \leqq n} w_i |f_i|, \tag{2.1}$$

where $f = \{f_i\}_{i=1}^n$, $w = \{w_i\}_{i=1}^n > 0$ and $\mathscr{M}$ consists of all h satisfying (1.4). For convenience denote the problem (1.2) for this case by P_∞ and the problem (1.5) by P_p, $1 \leqq p < \infty$. We investigate the convergence of the solution g_p of the problem P_p as $p \to \infty$.

Theorem 1. *Assume that there exists $w = \{w_i\}_{i=1}^n > 0$ such that*

$$0 < \liminf_{p\to\infty} (w_{p,i}/w_i^p) \leqq \limsup_{p\to\infty} (w_{p,i}/w_i^p) < \infty \tag{2.2}$$

for all i. Then the solution $g_p = \{g_{p,i}\}_{i=1}^n$ of the problem P_p converges as $p \to \infty$ to a solution $g_\infty = \{g_{\infty,i}\}_{i=1}^n$ of the problem P_∞ with weights $w = \{w_i\}_{i=1}^n$. Specifically

$$\begin{aligned} g_{\infty,i} &= \operatorname*{limit}_{p\to\infty} g_{p,i} = \max_{\{U: i\in U\}} \min_{\{L: i\in L\}} u_\infty(L \cap U) \\ &= \min_{\{L: i\in L\}} \max_{\{U: i\in U\}} u_\infty(L \cap U), \end{aligned} \tag{2.3}$$

for all i, where L and U are lower and upper subsets of X respectively and $u_\infty(L \cap U)$ is the unique real number satisfying

$$\max_{i\in L\cap U} w_i |f_i - u_\infty(L \cap U)| \leqq \max_{i\in L\cap U} w_i |f_i - u| \quad \text{for all } u. \tag{2.4}$$

Remarks. (i) If there exist δ, $m > 0$ such that $0 < \delta \leqq w_{p,i} \leqq m$ for all p and i, then (2.2) holds if and only if $w_i = 1$ for all i. In such a case the P_∞ problem has unit weights.

(ii) The solution of the problem P_p, $1 < p < \infty$ is unique and is given by (1.6). When $f = \{f_i\}_{i=1}^n$ is not isotone, the P_∞ problem has an infinite number of distinct solutions. This follows from the results in [11]. Theorem 1 indicates that exactly one of the solutions of the P_∞ problem is a limit of the solutions of the P_p problem when the hypothesis of Theorem 1 holds.

(iii) Compare (2.3) with the results in Ref. [11] of part I of this article.

Proof of Theorem 1. We first introduce some notation and prove two lemmas. For a fixed $f = \{f_i\}_{i=1}^n$ we define functions $\tau_p\colon R^n \to R$ and $\kappa_p\colon R \to R$ for $1 \leqq p \leqq \infty$ by

$$\tau_p(\mathbf{u}) = \sum_{i=1}^n w_{p,i} \, | f_i - u_i |^p, \qquad 1 \leqq p < \infty,$$

$$\tau_\infty(\mathbf{u}) = \max_{1 \leqq i \leqq n} w_i \, | f_i - u_i |,$$

$$\kappa_p(u) = \sum_{i=1}^n w_{p,i} \, | f_i - u |^p, \qquad 1 \leqq p < \infty,$$

$$\kappa_\infty(u) = \max_{1 \leqq i \leqq n} w_i \, | f_i - u |,$$

where $\mathbf{u} = (u_1, u_2, ..., u_n) \in R^n$ and $u \in R$.

LEMMA 1. *Assume* (2.2) *holds, then*

$$\operatorname*{limit}_{p \to \infty} (\tau_p(\mathbf{u}))^{1/p} = \tau_\infty(\mathbf{u}), \tag{2.5}$$

and

$$\operatorname*{limit}_{p \to \infty} (\kappa_p(u))^{1/p} = \kappa_\infty(u), \tag{2.6}$$

the convergence being uniform on compact sets in the domain of the respective function.

Proof. From (2.2) it follows that there exist real numbers $\delta_1, \delta_2 > 0$ and $p_0 \geqq 1$ such that $\delta_1 \leqq w_{p,i}/w_i{}^p \leqq \delta_2$ for all i for all $p \geqq p_0$. Hence

$$\tau_p(\mathbf{u}) = \sum_{i=1}^n (w_{p,i}/w_i{}^p)(w_i \, | f_i - u_i |)^p \leqq (n\delta_2)(\tau_\infty(\mathbf{u}))^p,$$

for all $p \geqq p_0$. Also by finiteness of X, there exists i_0 depending upon $\mathbf{u}$ such that $w_{i_0} \, | f_{i_0} - u_{i_0} | = \tau_\infty(\mathbf{u})$. Hence

$$\tau_p(\mathbf{u}) \geqq w_{p,i_0} \, | f_{i_0} - u_{i_0} |^p = (w_{p,i_0}/w_{i_0}^p)(w_{i_0} \, | f_{i_0} - u_{i_0} |)^p \geqq \delta_1(\tau_\infty(\mathbf{u}))^p.$$

It follows that

$$| \tau_p(\mathbf{u})^{1/p} - \tau_\infty(\mathbf{u}) | \leqq \max\{ |(n\delta_2)^{1/p} - 1 |, | \delta_1^{1/p} - 1 | \} \, \tau_\infty(\mathbf{u}).$$

Since $\tau_\infty(\mathbf{u})$ is continuous in $\mathbf{u}$, it is bounded on a compact domain. Hence the uniform convergence of $\tau_p(\mathbf{u})^{1/p}$ to $\tau_\infty(\mathbf{u})$ follows. Thus (2.5) is established. (2.6) follows from (2.5) with $\mathbf{u} = (u, u, ..., u)$.

Lemma 1 corresponds to the well-known result in measure theory concerning the function spaces $L_p(X, \Sigma, \mu)$, $1 \leqq p \leqq \infty$ that when $\mu(X) < \infty$, the L_p norm converges pointwise to the L_∞ norm as $p \to \infty$. See, e.g., [1]. Our setting owing to the introduction of the weight function w_p, which may vary with p, differs somewhat from the setting of the L_p spaces. Hence a condition such as (2.2) is required to prove convergence.

LEMMA 2. *Suppose for each p, $1 < p < \infty$ the real number u_p satisfies $\kappa_p(u_p) \leqq \kappa_p(u)$ for all u. Then u_p is unique, limit $u_p = u_\infty$ exists as $p \to \infty$ and $\kappa_\infty(u_\infty) \leqq \kappa_\infty(u)$ for all u. Further, κ_∞ has a unique minimizer.*

Proof. Note that $\kappa_p(u) \to \infty$ as $|u| \to \infty$ for $1 \leqq p \leqq \infty$. Since κ_p is continuous in u, by using compactness arguments we may show that a minimizer of κ_p exists for $1 \leqq p \leqq \infty$. It is easy to see that $\kappa_p(u)$ for $1 < p < \infty$ is strictly convex in u (see [5]). Hence, the minimizer u_p is unique. If v_∞ satisfies $\kappa_\infty(v_\infty) \leqq \kappa_\infty(u)$ for all u, then using the finiteness of X we have $\kappa_\infty(v_\infty) = w_{i_1}(x_{i_1} - v_\infty) = w_{i_2}(v_\infty - x_{i_2})$ for some i_1, i_2, $1 \leqq i_1$, $i_2 \leqq n$. It follows that the minimizer v_∞ of κ_∞ is also unique.

It is easy to see that

$$\min_{1 \leqq i \leqq n} x_i \leqq u_p \leqq \max_{1 \leqq i \leqq n} x_i ,$$

for all p, $1 < p < \infty$. Now let $t_k = u_{p_k}$, $k = 1, 2, \ldots$ be any subsequence of u_p such that $p_k \to \infty$ as $k \to \infty$. Since t_k are bounded there exists a convergent subsequence, say $t_{k_j} \to t_\infty$. We then have, letting $r_j = p_{k_j}$ for convenience,

$$(\kappa_{r_j}(t_{k_j}))^{1/r_j} \leqq (\kappa_{r_j}(u))^{1/r_j} \qquad \text{for all } u.$$

Since $\kappa_p^{1/p}$ is continuous and converges to κ_∞ uniformly on compact sets (Lemma 1), on letting $j \to \infty$ we have $\kappa_\infty(t_\infty) \leqq \kappa_\infty(u)$ for all u. It follows that $t_\infty = v_\infty$, since the minimizer is unique. Thus any subsequence u_{p_k} of u_p such that $p_k \to \infty$ as $k \to \infty$ contains in turn a subsequence converging to v_∞. Hence, limit $u_p = v_\infty$ as $p \to \infty$ and the assertions made in the lemma hold with $u_\infty = v_\infty$.

Now we proceed to the proof of Theorem 1. The solution $g_p = \{g_{p,i}\}_{i=1}^n$ of the problem P_p, $1 < p < \infty$ is given by (1.6). Considering $L \cap U$ instead of X in Lemma 2 we conclude that

$$\operatorname*{limit}_{p \to \infty} u_p(L \cap U) = u_\infty(L \cap U)$$

exists and (2.4) holds. Since the number of lower and upper sets is finite, from (1.6) it follows that the limit of $g_{p,i}$ exists as $p \to \infty$ for all i and (2.3) holds for some $g_{\infty,i}$. It now suffices to show that $\{g_{\infty,i}\}_{i=1}^n$ is a solution of the problem P_∞. Since $\{g_{p,i}\}_{i=1}^n$ is isotone for each p, $\{g_{\infty,i}\}_{i=1}^n$ also has this property. Clearly

$$\min_{1\leqq i\leqq n} x_i \leqq g_{p,i} \leqq \max_{1\leqq i\leqq n} x_i \qquad \text{for all } p, \qquad \text{all } i.$$

Using the definition of $\{g_{p,i}\}_{i=1}^n$ we have $\tau_p(\mathbf{g}_p)^{1/p} \leqq \tau_p(\mathbf{u})^{1/p}$ for all $\mathbf{u} \in R^n$. Letting $p \to \infty$ we conclude from (2.5) that $\tau_\infty(\mathbf{g}_\infty) \leqq \tau_\infty(\mathbf{u})$ for all $\mathbf{u} \in R^n$. Hence $\{g_{\infty,i}\}_{i=1}^n$ is a solution of the problem P_∞. The proof of Theorem 1 is now complete.

3. Norm Reducing Property of a Solution of the Isotone Optimization Problem with Respect to the Uniform Norm

Let X be an arbitrary partially ordered set and $w(x) = 1$ for all $x \in X$. Then the norm (1.1) becomes the uniform norm $\|\cdot\|$, where

$$\|f\| = \sup_{x\in X} |f(x)|, \qquad f \in \mathscr{V}. \tag{3.1}$$

We consider the problem (1.2) with $\|\cdot\|_w$ replaced by $\|\cdot\|$. For $f \in \mathscr{V}$ define $f^0 \in \mathscr{V}$ by

$$f^0 = (1/2)(\sup_{\{z:z\leq x\}} f(z) + \inf_{\{z:x\leq z\}} f(z)), \qquad x \in X. \tag{3.2}$$

It is easy to see that f^0 is isotone.

Theorem 2. *Let $f, f_1, f_2 \in \mathscr{V}$. Then*

(i) $$\|f - f^0\| = \min_{h\in\mathscr{M}} \|f - h\|, \tag{3.3}$$

i.e., f^0 solves the problem (1.2) *for the norm* $\|\cdot\|$.

(ii) $$\|f_1^0 - f_2^0\| \leqq \|f_1 - f_2\|, \tag{3.4}$$

i.e., the norm reducing property holds.

Proof.

(i) Let

$$\theta = (1/2) \sup_{\{(x,y)\in X\times X: x\leq y\}} (f(x) - f(y)), \tag{3.5}$$

$$\underline{g}(x) = \sup_{\{z:z\leq x\}} f(z) - \theta, \qquad x \in X, \tag{3.6}$$

$$\bar{g}(x) = \inf_{\{z:x\leq z\}} f(z) + \theta, \qquad x \in X, \tag{3.7}$$

then

$$f^0 = (1/2)(\underline{g} + \bar{g}). \tag{3.8}$$

We know from the results of Section 2 of [11] that

$$\theta = \min_{h \in \mathscr{M}} \| f - h \| = \| f - \underline{g} \| = \| f - \bar{g} \|. \tag{3.9}$$

But since $\underline{g} \leqq f^0 \leqq \bar{g}$, (i) again follows from the results in Section 2 of [11].

(ii) Let θ_i, $\underline{g}_i$, $\bar{g}_i$, $i = 1, 2$ be defined by (3.5), (3.6), (3.7), respectively, with f_i, $i = 1, 2$, in the right-hand sides of these expressions. Again (3.8), (3.9) hold with $\theta, f, f^0, \underline{g}$ and $\bar{g}$ replaced respectively by θ_i, f_i, f_i^0, $\underline{g}_i$ and $\bar{g}_i$ for each $i = 1, 2$. Let $x \in X$ and $\epsilon > 0$. Then by the definition of $\underline{g}_i$, $\bar{g}_i$ there exist z_1, $z_2 \in X$ such that $z_1 \leq x \leq z_2$ and

$$\underline{g}_1(x) \leqq f_1(z_1) - \theta_1 + \epsilon,$$
$$\bar{g}_2(x) \geqq f_2(z_2) + \theta_2 - \epsilon.$$

Also

$$\bar{g}_1(x) \leqq f_1(z_2) + \theta_1,$$
$$\underline{g}_2(x) \geqq f_2(z_1) - \theta_2.$$

Using (3.8) we may derive from the above four inequalities the following:

$$f_1^0(x) - f_2^0(x) \leqq (1/2)(f_1(z_1) - f_2(z_1)) + (1/2)(f_1(z_2) - f_2(z_2)) + \epsilon.$$

Hence

$$f_1^0(x) - f_2^0(x) \leqq \| f_1 - f_2 \| + \epsilon.$$

Interchanging subscripts 1 and 2 and noting that ϵ, x are arbitrary we conclude that (3.4) holds. The proof is now complete.

Remarks. According to the results of Section 2 of [11] any g in $\mathscr{M}$ satisfying $\underline{g} \leqq g \leqq \bar{g}$ minimizes $\| f - h \|$ for h in $\mathscr{M}$. We have indeed isolated an f^0 in $\mathscr{M}$ from this infinite set of minimizers such that (3.4) holds. It is shown in the Ref. [8] of part I of this article that a similar result is true under certain conditions for the function space $L_\infty(X, \Sigma, \mu)$, where X is a totally ordered set. The result also holds for the L_2 norm case. See Dykstra [4].

4. Differentiability Properties of $\underline{g}$ and $\bar{g}$

We now consider the problem (1.2) with $X = [a, b]$, a closed interval of the real line. We showed in [11] that both $\underline{g}$ and $\bar{g}$ solve the problem (1.2). Here

$$\underline{g}(x) = \sup_{z \in [a,x]} (f(z) - \theta/w(z)), \qquad x \in [a, b], \tag{4.1}$$

$$\bar{g}(x) = \inf_{z \in [x,b]} (f(z) + \theta/w(z)), \qquad x \in [a, b], \tag{4.2}$$

and

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

where

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

In this section we investigate the differentiability and other properties of g and $\bar{g}$.

We first introduce concepts called the Level and Descent Sets. Let $f \in \mathscr{V}$ and define for each $x \in [a, b]$ the following sets:

$$L(f, x) = \bigcup \{[x, y]: x < y \leqq b \text{ and } f(z) = f(x) \text{ for all } z \in [x, y]\},$$
$$D_1(f, x) = \bigcup \{(x, y): x < y \leqq b \text{ and } f(z) < f(x) \text{ for all } z \in (x, y)\},$$
$$D_2(f, x) = \bigcup \{[x, y]: x < y \leqq b \text{ and } f(z) \leqq f(x) \text{ for all } z \in [x, y]\}.$$

Define the Level Set $L(f)$ and the Strong Descent Set $D_1(f)$ by

$$L(f) = \bigcup_{x\in[a,b]} L(f, x),$$

$$D_1(f) = \bigcup_{x\in[a,b]} D_1(f, x).$$

Also define the Weak Descent Set $D_2(f)$ by

$$D_2(f) = \bigcup_{x\in[a,b]} D_2(f, x).$$

We now state

LEMMA 3. *Let $f \in \mathscr{C}$.*

(i) *If $L(f) \neq \varnothing$ then*

$$L(f) = \bigcup_{n=1}^{\infty} [\alpha_n, \beta_n], \qquad a \leqq \alpha_n < \beta_n \leqq b \qquad \textit{for all } n,$$

where $[\alpha_n, \beta_n]$ are disjoint closed intervals such that $f(z) = f(\alpha_n)$ for all $z \in [\alpha_n, \beta_n]$.

(ii) *If $D_1(f) \neq \varnothing$ then*

$$D_1(f) = \bigcup_{n=1}^{\infty} (\rho_n, \sigma_n), \qquad a \leqq \rho_n < \sigma_n \leqq b \qquad \textit{for all } n,$$

where (ρ_n, σ_n) are disjoint open intervals such that $f(z) < f(\rho_n)$ for all $z \in (\rho_n, \sigma_n)$ and $f(x) \leqq f(\rho_n)$ for all $x \in [a, \rho_n]$.

(iii) *If $D_2(f) \neq \varnothing$ then*

$$D_2(f) = \bigcup_{n=1}^{\infty} [\lambda_n, \mu_n], \qquad a \leqq \lambda_n < \mu_n \leqq b \qquad \text{for all } n,$$

where $[\lambda_n, \mu_n]$ are disjoint closed intervals such that $f(z) \leqq f(\lambda_n)$ for all $z \in [\lambda_n, \mu_n]$ and $f(x) < f(\lambda_n)$ for all $x \in [a, \lambda_n)$.

Clearly, whenever $L(f) \neq \varnothing$, $D_2(f) \neq \varnothing$ we have,

$$L^{(0)}(f) = \text{interior of } L(f) = \bigcup_{n=1}^{\infty} (\alpha_n, \beta_n),$$

$$D_2^{(0)}(f) = \text{interior of } D_2(f) = \bigcup_{n=1}^{\infty} (\lambda_n, \mu_n).$$

We denote by $f^{(k)}(x)$, the kth derivative of f at x, if it exists. We define P as in Section 2 of [11] and recall that if f and w are continuous then so is g.

THEOREM 3. *Let μ^* be the Lebesgue measure on $[a, b]$.*

(A) *Let $f, w \in \mathscr{C}$ then*

(i) $(f \in \mathscr{M}) \Leftrightarrow (D_1(f) = \varnothing) \Leftrightarrow (D_2(f) = L(f))$

(ii) $L(g) = D_2(f - \theta/w) \supset P$

(iii) $\{x \in [a, b]: g(x) = f(x) - \theta/w(x)\} = [a, b] - D_1(f - \theta/w)$

(iv) $\{x \in [a, b]: g^{(n)}(x) = 0, n = 1, 2,...\} \supset D_2^{(0)}(f - \theta/w)$

(v) *$g^{(1)}(x)$ may not exist at most on a set*

$$E \subset [a, b] - D_2^{(0)}(f - \theta/w)$$

with $\mu^*(E) = 0$.

(B) *$\mu^*(D_2(f)) = \mu^*(D_2^{(0)}(f))$ for all $f \in \mathscr{C}$.*

If $f, w \in \mathscr{C}$, $g \not\equiv$ constant and

$$\mu^*(D_2(f - \theta/w) = \mu^*(D_2^{(0)}(f - \theta/w)) = b - a,$$

which, of course, implies from A(iv) *that $g^{(n)}(x) = 0$, μ^* —* a.e. *on $[a, b]$ for $n = 1, 2,...$, then both f and w cannot be absolutely continuous.*

Remarks.

(i) A similar theorem may be stated for $\bar{g}$, but in this case we need to modify the definitions of L, D_1 and D_2 sets. For example, we may define the set D_1' as follows

$$D_1'(f, x) = \bigcup \{(y, x): a \leqq y < x \quad \text{and} \quad f(z) > f(x) \text{ for all } z \in (y, x)\}, \qquad x \in [a, b]$$

$$D_1'(f) = \bigcup_{x \in [a,b]} D_1'(f, x).$$

Similar modifications necessary for the definitions of other sets are evident.

(ii) We give below one example of special interest to illustrate the results of Theorem 3. It will be seen that $g^{(1)}(x)$ does not exist on the set $E = [a, b] - D_2^{(0)}(f - \theta/w)$ with $\mu^*(E) = 0$. (See Theorem 3, A(v).) Let $[a, b] = [0, 1]$ and $f: [0, 1] \to [0, 1]$ be the well-known Cantor ternary function, ([6], p. 138). Then f is nondecreasing continuous with range [0, 1]. Hence $\theta = 0$ and $g = \bar{g} = f$. Let K be the Cantor ternary set. Then $[0, 1] - K$ is the union of disjoint open intervals and f is constant on each of these intervals. Clearly $D_2^{(0)}(f - \theta/w) = D_2^{(0)}(f) = [0, 1] - K$ and on $[0, 1] - K$, $f^{(1)}(x)$ exists and equals 0. $f^{(1)}$ does not exist on $E = K$. It is a known fact that $\mu^*(E) = 0$.

We prove Lemma 3 before proceeding to the proof of Theorem 3.

Proof of Lemma 3. We prove (iii). The proofs for (i) and (ii) are similar. Let $t \in D_2(f)$. Define

$$\lambda_t = \inf\{x: t \in [x, y], a \leqq x < y \leqq b \text{ and } f(z) \leqq f(x) \text{ for all } z \in [x, y]\},$$
$$\mu_t = \sup\{y: t \in [x, y], a \leqq x < y \leqq b \text{ and } f(z) \leqq f(x) \text{ for all } z \in [x, y]\}.$$

There exist $[x_n, y_n]$, $n = 1, 2,\ldots$ such that $t \in [x_n, y_n]$, $a \leqq x_n < y_n \leqq b$, $f(z) \leqq f(x_n)$ for all $z \in [x_n, y_n]$ and $x_n \to \lambda_t$ as $n \to \infty$. We may take $x_{n+1} \leqq x_n \cdots \leqq x_1 < y_1$. Since $x_n \leqq x_1 \leqq t \leqq y_1$ we have $f(z) \leqq f(x_1)$ for all $z \in [x_1, y_1]$ and $f(z) \leqq f(x_n)$ for all $z \in [x_n, t]$. Hence $f(z) \leqq f(x_n)$ for all $z \in [x_n, y_1]$. Since $f(x_{n+1}) \geqq (f(x_n))$, using continuity of f we have $f(\lambda_t) \geqq f(z)$ for all $z \in [\lambda_t, y_1]$. Thus, $\lambda_t \in D_2(f)$. Suppose $t \in [x, y]$, where $a \leqq x < y \leqq b$ and $f(z) \leqq f(x)$ for all $z \in [x, y]$, then $x \in [\lambda_t, y]$ and $f(\lambda_t) \geqq f(z)$ for all $z \in [\lambda_t, y]$. Hence

$$\mu_t = \sup\{y: \lambda_t < y \leqq b \text{ and } f(z) \leqq f(\lambda_t) \text{ for all } z \in [\lambda_t, y]\}.$$

Clearly, $\lambda_t < \mu_t$ and $f(z) \leqq f(\lambda_t)$ for all $z \in [\lambda_t, \mu_t]$. Thus $[\lambda_t, \mu_t] \subset D_2(f)$. Hence,

$$D_2(f) = \bigcup_{t \in D_2(f)} [\lambda_t, \mu_t].$$

Suppose $u \in [\lambda_t, \mu_t]$ then $\lambda_u \leqq \lambda_t$ and $\mu_u \geqq \mu_t$. Since $t \in [\lambda_t, \mu_t]$ implies $t \in [\lambda_u, \mu_u]$ we have $\lambda_t \leqq \lambda_u$ and $\mu_t \geqq \mu_u$. Thus, $[\lambda_t, \mu_t] = [\lambda_u, \mu_u]$. The intervals are therefore disjoint and the countability follows since each of the intervals includes a distinct rational number.

Suppose now there exists x, $a \leqq x < \lambda_n$ for some n such that $f(x) \geqq f(\lambda_n)$. Define

$$v = \inf\{u \in [x, \lambda_n]: f(u) = \max_{z \in [x, \lambda_n]} f(z)\}.$$

Then $a \leqq v < \lambda_n$ and $f(z) \leqq f(v)$ for all $z \in [v, \lambda_n]$ which is a contradiction to the definition of λ_n.

LEMMA 4.

(i) $(f \in \mathscr{M}) \Rightarrow (D_1(f) = \varnothing) \Leftarrow (D_2(f) = L(f))$.

(ii) *Suppose* $f \in \mathscr{C}$ *then*

$$(f \in \mathscr{M}) \Leftrightarrow (D_1(f) = \varnothing) \Leftrightarrow (D_2(f) = L(f)).$$

Proof of the lemma is simple. Note that if $f \in \mathscr{V} - \mathscr{C}$ then $D_1(f) = \varnothing$ does not imply that $f \in \mathscr{M}$ or $D_2(f) = L(f)$. As an example take f: $[0, 1] \to R$ defined by $f(x) = 1$, $x \in [0, 1)$ and $f(1) = 0$. Similarly $D_2(f) = L(f)$ does not imply that $f \in \mathscr{M}$. Take for example f: $[0, 1] \to R$ given by $f(x) = 0$, $x \in [0, 1/4] \cup \{1/2\}$, $f(x) = x$ otherwise.

Proof of Theorem 3.

A(i) This is part (ii) of Lemma 4.

(ii) Let $u \in L(g) = \bigcup_{n=1}^{\infty} [\xi_n, \eta_n]$ by Lemma 3, then $u \in [\xi_n, \eta_n]$ some n and $g(x) = g(\xi_n)$ for all $x \in [\xi_n, \eta_n]$. Since $g \in \mathscr{M}$, using the properties of $L(g)$ we conclude that $g(x) < g(\xi_n)$ for all $x \in [a, \xi_n)$. It follows from the definition of g that $g(\xi_n) = f(\xi_n) - \theta/w(\xi_n)$ and

$$f(x) - \theta/w(x) \leqq f(\xi_n) - \theta/w(\xi_n) \qquad \text{for all} \quad x \in [\xi_n, \eta_n].$$

It follows that $u \in L(f - \theta/w)$.

Now if $u \in D_2(f - \theta/w) = \bigcup_{n=1}^{\infty} [\gamma_n, \delta_n]$ by Lemma 3, then $u \in [\gamma_n, \delta_n]$ some n. Also

$$f(x) - \theta/w(x) \leqq f(\gamma_n) - \theta/w(\gamma_n) \qquad \text{for all} \quad x \in [\gamma_n, \delta_n].$$

Using the definition of g we conclude that $g(x) = g(\gamma_n)$ for all $x \in [\gamma_n, \delta_n]$. Hence $u \in L(g)$. This proves the equality of two sets. The assertion concerning P follows from the properties of P established in Section 2 of [11].

(iii) Let $x \in [a, b] - D_1(f - \theta/w)$. We assert that if $a \leqq y < x$ then

$$f(y) - \theta/w(y) \leqq f(x) - \theta/w(x).$$

If, on the contrary, for some y, $a \leqq y < x$,

$$f(y) - \theta/w(y) > f(x) - \theta/w(x)$$

holds, then let

$$t = \sup\{u\colon u \in [y, x], f(u) - \theta/w(u) = \max_{z \in [y,x]} (f(z) - \theta/w(z))\}.$$

Clearly $a \leqq t < x$ and by continuity of $f - \theta/w$ there exists v, $x < v$ such that

$$f(z) - \theta/w(z) < f(t) - \theta/w(t) \qquad \text{for all} \quad z \in (t, v).$$

Hence, $x \in D_1(f - \theta/w)$, a contradiction. This establishes the validity of the assertion made above. It follows from the definition of g that $g(x) = f(x) - \theta/w(x)$.

Now suppose $x \in D_1(f - \theta/w) = \bigcup_{n=1}^{\infty} (\gamma_n, \delta_n)$ by Lemma 3. Then $x \in (\gamma_n, \delta_n)$ some n. Hence

$$f(z) - \theta/w(z) < f(\gamma_n) - \theta/w(\gamma_n), \qquad \text{for all} \quad z \in (\gamma_n, \delta_n).$$

We then have

$$g(x) \geqq g(\gamma_n) \geqq f(\gamma_n) - \theta/w(\gamma_n) > f(x) - \theta/w(x).$$

(iv) From Lemma 3 it follows that $L(g) = \bigcup_{n=1}^{\infty} [\xi_n, \eta_n]$, where $[\xi_n, \eta_n]$ are disjoint intervals. Hence $L^{(0)}(g) = \bigcup_{n=1}^{\infty} (\xi_n, \eta_n)$ is an open set and $g^{(n)}(x) = 0$ for all $x \in L^{(0)}(g)$, $n = 1, 2,\dots$. From part (ii) we have $L^{(0)}(g) = D_2^{(0)}(f - \theta/w)$.

(v) Since g is nondecreasing, it follows that g is differentiable μ^* — a.e. (see [9], p. 96.) By (iv), $g^{(1)}(x) = 0$ on $D_2^{(0)}(f - \theta/w)$ and the result follows.

(B) If $D_2(f) = \varnothing$ then $D_2^{(0)}(f) = \varnothing$ and the μ^*-measures of these two sets are equal to 0. Suppose $D_2(f) \neq \varnothing$, then by Lemma 3 $D_2(f) = \bigcup_{n=1}^{\infty} [\lambda_n, \mu_n]$ where the intervals $[\lambda_n, \mu_n]$ all disjoint. Hence $D_2^{(0)}(f) = \bigcup_{n=1}^{\infty} (\lambda_n, \mu_n)$. It follows that the set $D_2(f) - D_2^{(0)}(f)$ is at most countable and therefore the μ^*-measures of these sets are equal.

The proof of the remaining part is similar to the one used in showing that the Cantor ternary function is not absolutely continuous. Let

$$H = [a, b] - D_2^{(0)}(f - \theta/w).$$

Clearly H is compact and $a, b \in H$. By hypothesis $\mu^*(H) = 0$. Let $\epsilon > 0$. From the theory of Lebesgue measure we conclude that there exists a countable sequence of open intervals (x_i, y_i), $i = 1, 2,..., -\infty < x_i < y_i < \infty$ such that $H \subset \bigcup_{i=1}^{\infty} (x_i, y_i)$ and $\sum_{i=1}^{\infty} | y_i - x_i | < \epsilon$. Since H is compact, by taking finite unions and renumbering if necessary, we can find a finite covering (x_i, y_i), $i = 1, 2,..., n$ of H such that $x_i < y_i < x_{i+1} < y_{i+1}$, $i = 1, 2,..., n-1$ and $a \in (x_1, y_1)$, $b \in (x_n, y_n)$. Now $D_2^{(0)}(f - \theta/w) \subset L(g)$ by A(ii) and the former set is a countable union of open intervals on each of which g is constant, it follows that $g(y_i) = g(x_{i+1})$, $i = 1, 2,..., n-1$. Thus

$$| b - x_n | + \sum_{i=2}^{n-1} | y_i - x_i | + | y_1 - a | < \epsilon$$

and

$$\eta = g(b) - g(a) = g(b) - g(x_n) + \sum_{i=2}^{n-1} (g(y_i) - g(x_i)) + g(y_i) - g(a)$$

Since by hypothesis $\eta > 0$, it follows that g is not absolutely continuous. We showed in Theorem 2 of [11] that if f and w are absolutely continuous then so is g. Hence both f and w cannot be absolutely continuous.

The proof of Theorem 3 is now complete.

We remark that part B may also be proved by applying Theorem 13, p. 106 of [9] to g.

5. Algorithms

In this section we consider the problem (1.2) with $X = [a, b]$ as in Section 4 and develop algorithms to compute g and $\bar{g}$ defined by (4.1) and (4.2), respectively. Specifically we let G_n, $n = 1, 2,...$ be a sequence of finite sets contained in $[a, b]$ such that G_n becomes dense in $[a, b]$ as $n \to \infty$ and construct a sequence of functions $g_n(\bar{g}_n)$, $n = 1, 2,...$ defined on $[a, b]$ but depending on G_n such that $g_n(\bar{g}_n)$ converges uniformly to $g(\bar{g})$ as $n \to \infty$. We also establish rates of convergence of various quantities involved. We now state the following:

THEOREM 4. *Let $f, w \in \mathscr{C}, f \notin \mathscr{M}$. Let $G_n \subset [a, b]$, $n = 1, 2,...$ be a sequence of finite sets such that $a, b \in G_n$ for all n and*

$$\delta_n = \sup_{x \in [a,b]} \inf_{y \in G_n} | x - y | \to 0 \quad \text{as} \quad n \to \infty.$$

Let

$$S_n = \{(x, y) \in G_n \times G_n \colon x, y \in G_n\,, x \leqq y\}$$

and

$$\theta_n = \max_{(x,y)\in S_n} \frac{w(x)\,w(y)}{w(x) + w(y)} (f(x) - f(y)).$$

Define also

$$M_f = \max_{x\in[a,b]} |f(x)|,$$

$$M_w = \max_{x\in[a,b]} |w(x)|,$$

$$m_w = \min_{x\in[a,b]} |w(x)|.$$

We then have

(A) *$\theta_n \leqq \theta$ for all n and $\theta_n \to \theta$ as $n \to \infty$ according to*

$$0 \leqq \theta - \theta_n \leqq (M_w{}^2/m_w)\,\lambda(f, \delta_n) + (M_w{}^2 M_f/m_w{}^2)\,\lambda(w, \delta_n), \tag{5.1}$$

where

$$\lambda(h, \delta) = \max_{|x-y|\leqslant\delta, x,y\in[a,b]} |h(x) - h(y)|, \qquad h \in \mathscr{C}$$

is the modulus of continuity of h. (*See* [8].)

(B) *Define $\underline{g}_n$, $\bar{g}_n$: $[a, b] \to R$ by*

$$\underline{g}_n(x) = \max_{z\in[a,x]\cap G_n} (f(z) - \theta_n/w(z)), \qquad x \in G_n\,,$$

$$\bar{g}_n(x) = \min_{z\in[x,b]\cap G_n} (f(z) + \theta_n/w(z)), \qquad x \in G_n\,,$$

and for $x \in [a, b] - G_n$ choose any value of $\underline{g}_n(x)(\bar{g}_n(x))$ that will make the function $\underline{g}_n(\bar{g}_n)$ nondecreasing on $[a, b]$ (e.g., choose linear interpolation or form a step function). Then $g_n(= \underline{g}_n$ or $\bar{g}_n)$ converges to $g(= \underline{g}$ or $\bar{g}$, respectively) uniformly according to

$$\sup_{x\in[a,b]} |g(x) - g_n(x)|$$
$$\leqq M_w{}^2/m_w{}^2 + 2)\,\lambda(f, \delta_n) + (1/m_w{}^2)(M_w{}^2 M_f/m_w + 2\theta)\,\lambda(w, \delta_n). \tag{5.2}$$

(C) *Define $\underline{g}_n$, $\bar{g}_n$: $[a, b] \to R$ by*

$$\underline{g}_n(x) = \max_{z\in[a,x]\cap G_n} (f(z) - \theta/w(z)), \qquad x \in G_n\,,$$
$$\bar{g}_n(x) = \min_{z\in[x,b]\cap G_n} (f(z) + \theta/w(z)), \qquad x \in G_n\,,$$

and for $x \in [a, b] - G_n$ *choose any value of* $g_n(x)(\bar{g}_n(x))$ *as in* (B) *that will make the function* $g_n(\bar{g}_n)$ *nondecreasing on* $[a, b]$. *Then* $g_n \leqq g \leqq \bar{g} \leqq \bar{g}_n$ *and* $g_n(= g_n$ *or* $\bar{g}_n)$ *converges to* $g(= g$ *or* $\bar{g}$, *respectively) uniformly according to*

$$\sup_{x \in [a,b]} | g(x) - g_n(x)| \leqq 2\lambda(f, \delta_n) + (2\theta/m_w^2)\, \lambda(w, \delta_n). \tag{5.3}$$

(D) *If* G_n *in addition satisfies*

$$G_n \subset G_{n+1} \qquad \text{for all } n,$$

then

$$g_n(x) \leqq g_{n+m}(x) \leqq \bar{g}_{n+m}(x) \leqq \bar{g}_n(x) \qquad \text{for all} \quad x \in G_n \quad \text{all } n, m \geqq 1,$$

where g_n *and* $\bar{g}_n$ *are as defined in* (B) *or* (C).

Proof of Theorem 4.

(A) Clearly, $\theta_n \leqq \theta$. There exist $x, y \in [a, b]$, $x < y$ and $f(x) > f(y)$ such that

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

It is easy to see that there exist $x_n, y_n \in G_n$ such that $| x - x_n | \leqq \delta_n$, $| y - y_n | \leqq \delta_n$, $x_n \leqq y_n$. Then

$$\theta_n \geqq (w(x_n)\, w(y_n)/(w(x_n) + w(y_n)))(f(x_n) - f(y_n)).$$

Hence,

$$\theta - \theta_n \leqq \frac{w(x)\, w(y)}{w(x) + w(y)} (f(x) - f(x_n) + f(y_n) - f(y)) + \left(\frac{w(x)\, w(y)}{w(x) + w(y)} - \frac{w(x_n)\, w(y_n)}{w(x_n) + w(y_n)}\right)(f(x_n) - f(y_n)). \tag{5.4}$$

Now

$$\frac{w(x)\, w(y)}{w(x) + w(y)} - \frac{w(x_n) w(y_n)}{w(x_n) + w(y_n)} = \frac{w(x)\, w(x_n)(w(y) - w(y_n)) + w(y)\, w(y_n)(w(x) - w(x_n))}{(w(x) + w(y))(w(x_n) + w(y_n))}.$$

Hence

$$\frac{w(x)\, w(y)}{w(x) + w(y)} - \frac{w(x_n)\, w(y_n)}{w(x_n) + w(y_n)} \leqq \frac{M_w^2}{2m_w^2} \lambda(w, \delta_n).$$

Also $| f(x_n) - f(y_n)| \leqq 2M_f$ and

$$\frac{w(x)\, w(y)}{w(x) + w(y)} (f(x) - f(x_n) + f(y_n) - f(y)) \leqq \frac{M_w^2}{m_w} \lambda(f, \delta_n).$$

Using the above bounds in (5.4) we may deduce (5.1).

(B) We show the result for g_n, the proof for $\bar{g}_n$ is similar. Let $x \in [a, b]$, then by the definition of g, there exists $z \in [a, x]$ such that $g(x) = f(z) - \theta/w(z)$. There exists $u \in G_n$ such that $0 \leqq z - u \leqq 2\delta_n$ and hence

$$g_n(x) \geqq g_n(u) \geqq f(u) - \theta_n/w(u) \geqq f(u) - \theta/w(u),$$

the last inequality following from the fact that $\theta_n \leqq \theta$. We conclude that

$$\begin{aligned} g(x) - g_n(x) &\leqq f(z) - f(u) + \theta(w(z) - w(u))/(w(z)\, w(u)) \\ &\leqq \lambda(f, 2\delta_n) + (\theta/m_w^2)\, \lambda(w, 2\delta_n). \end{aligned}$$

Since $\lambda(f, k\delta_n) \leqq k\lambda(f, \delta_n)$, where k is a positive integer, we have,

$$g(x) - g_n(x) \leqq 2\lambda(f, \delta_n) + (2\theta/m_w^2)\, \lambda(w, \delta_n). \tag{5.5}$$

Now if $x \in [a, b]$, then there exists $v \in [a, b]$ such that $0 \leqq v - x \leqq 2\delta_n$. Then by the definition of g_n, we have

$$g_n(x) \leqq g_n(v) = f(t) - \theta_n/w(t) \tag{5.6}$$

for some $t \in [a, v] \cap G_n$. If $t \leqq x$, we have

$$g(x) \geqq g(t) \geqq f(t) - \theta/w(t)$$

and from (5.6), (5.1) we conclude that

$$g_n(x) - g(x) \leqq (\theta - \theta_n)/w(t) \leqq (M_w^2/m_w^2)\, \lambda(f, \delta_n) + (M_w^2 M_f/m_w^3)\, \lambda(w, \delta_n). \tag{5.7}$$

If, on the other hand, $x < t \leqq v$, we observe that $g(x) \geqq f(x) - \theta/w(x)$ and from (5.6) obtain,

$$g_n(x) - g(x) \leqq f(t) - f(x) + (\theta - \theta_n)/w(t) + \theta(w(t) - w(x))/(w(t)\, w(x))$$

which by (5.1) reduces to

$$g_n(x) - g(x) \leqq (M_w^2/m_w^2 + 2)\, \lambda(f, \delta_n) + (1/m_w^2)(M_w^2 M_f/m_w + 2\theta)\, \lambda(w, \delta_n). \tag{5.8}$$

Comparing (5.7) and (5.8) we see that (5.8) holds for all $x \in [a, b]$. The required result (5.2) is then derived from (5.5) and (5.8).

(C) This may be proved using arguments similar to those used to prove (B).

(D) This is evident.

References

1. R. G. Bartle, "The Elements of Integration," John Wiley and Sons, Inc., New York, 1966.
2. H. D. Brunk, Conditional expectation given a σ-lattice and applications, *Ann. Math. Statist.* **36** (1965), 1339–1350.
3. H. D. Brunk and S. Johansen, A generalized Radon–Nikodym derivative, *Pacific J. Math.* **34** (1970), 585–617.
4. R. L. Dykstra, A characterization of a conditional expectation with respect to a σ-lattice, *Ann. Math. Statist.* **41** (1970), 698–701.
5. O. L. Mangasarian, "Nonlinear Programming," McGraw–Hill, New York, 1969.
6. M. E. Munroe, "Measure and Integration," 2nd. ed., Addison–Wesley, Reading, MA, 1971.
7. W. T. Reid, A simple optimal control problem involving approximation by monotone functions, *J. Optimization Theory Appl.* **2** (1968), 365–377.
8. J. R. Rice, "The approximation of functions," Vol. I, Addison–Wesley, Reading, MA, 1964.
9. H. L. Royden, "Real Analysis," 2nd ed., Macmillan Company, New York, 1968.
10. 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 12B AD-734 130, February 10, 1972.
11. V. A. Ubhaya, Isotone optimization. I, *J. Approximation Theory*, **12** (1974), 146–159.
12. C. Van Eeden, Testing and estimating ordered parameters of probability distributions, Ph.D. dissertation, University of Amsterdam, Studentendrukkerij Poortpers N. V., Amsterdam, 1958.