JOURNAL OF APPROXIMATION THEORY 36, 304–316 (1982)

Least-Squares Approximation with Restricted Range

KENNETH M. LEVASSEUR

Department of Mathematics, University of Lowell, Lowell, Massachusetts 01854, USA

AND

JAMES T. LEWIS

Department of Mathematics, University of Rhode Island, Kingston, Rhode Island 02881, USA

Communicated by Oved Shisha

Received October 23, 1981

1. INTRODUCTION

We will consider the least-squares restricted range approximation problem

$$\underset{h \in H}{\text{minimize}} \int_{A_2} w(f-h)^2 \text{ subject to } v(t) \leqslant h(t) \leqslant u(t) \text{ for all } t \text{ in } A_1.$$

Here f, w, v, and u are given functions and H is a family of approximating functions; precise hypotheses will be specified in the next section. Restricted range approximation includes as special cases positive approximation (where $h(t) \geqslant 0$ is required) and one-sided approximation (where for approximation from below, say, $h(t) \leqslant f(t)$ is required). G. D. Taylor and others have studied uniform approximation with restricted range, where $\max_x |f(x)-h(x)|$ is minimized, subject to the same constraints as above, and have developed a complete theory, cf. [13].

Least-squares approximation by positive polynomials has been considered in [1] and [10]. In [9] a problem of digital filter design has been formulated as a problem of least-squares approximation with restricted range.

Our main result in this paper is a characterization theorem for least-squares approximation with restricted range. We also obtain several corollaries and state the generalization of the characterization theorem to the problem of L_{2p} approximation with restricted range, where $2p$ is an even integer. In a subsequent paper we will discuss computational algorithms and

0021-9045/82/120304-13$02.00/0

numerical examples for least-squares approximation with restricted range. We have found that the conditions specified to characterize a best approximation can be conveniently verified in the examples we have considered.

Gehner [4, 5] has used an optimization theory approach to obtain a general characterization theorem [5, Theorem 4, p. 54]). This theorem was used in [5] to obtain characterization results for constrained Chebyshev and L_1 approximation. This approach could also be employed to obtain characterization results for constrained L_p approximation, $1 < p < \infty$.

After completing this paper we learned of the work of Evans and Cantoni [3], which contains, among other results, two characterization theorems for the least-squares restricted range problem.

2. The Characterization Theorem

In this section we list our hypotheses, note the existence of a unique solution to the least-squares restricted range problem, and then prove a characterization theorem giving necessary and sufficient conditions for an approximation to be the best (least squares) approximation with restricted range. First the hypotheses:

H_1: A_1 is a closed and bounded set of real numbers and A_2 is a finite union of closed, bounded intervals of real numbers.

H_2: f is a real continuous function on $A_1 \cup A_2$.

H_3: v and u are real continuous functions on A_1 with $v(t) < u(t)$ for all t in A_1.

H_4: $\{h_1, ..., h_n\}$ is a set of real continuous functions on $A_1 \cup A_2$ which is linearly independent on A_2; denote by H the set of all linear combinations (with real coefficients) of $h_1, ..., h_n$.

H_5: w is a real, continuous strictly positive (weight) function on A_2.

H_6: The exists a linear combination $h = \sum_{i=1}^{n} a_i h_i$ such that $v(t) \leqslant h(t) \leqslant u(t)$ for all t in A_1.

H_7: For $k = 1, ..., n$ the set $\{h_1, ..., h_k\}$ is a Chebyshev system of order k on A_1; that is, if $t_1, ..., t_k$ are distinct points in A_1, then $\det(h_j(t_i)) \neq 0$.

For a subset $B \subseteq A_1$, the *least-squares restricted range problem* will be denoted by R_B and is:

$$\underset{h \in H}{\text{minimize}} \int_{A_2} w(f-h)^2 \text{ subject to } v(t) \leqslant h(t) \leqslant u(t) \text{ for all } t \text{ in } B.$$

Of course h^* in H is a *solution* of R_B if $v(t) \leqslant h^*(t) \leqslant u(t)$ for all t in B and if

$$\int_{A_2} w(f-h^*)^2 \leqslant \int_{A_2} w(f-h)^2 \qquad \text{for every } h \text{ in } H$$

which satisfies $v(t) \leqslant h(t) \leqslant u(t)$ for all t in B.

When $B = \varnothing$, the problem reduces to the familiar unconstrained least-squares problem.

Since the following existence and uniqueness result can be established using standard arguments, its proof will be omitted; cf. [11].

THEOREM 1. *If* H_1–H_6 *are satisfied and* $B \subseteq A_1$, *then the least-squares restricted range problem* R_B *has a unique solution.*

We now proceed to develop the characterization theorem. The basic idea is this: known results on convex programming give a characterization theorem when A_1 has a finite number of points; then a discretization result yields the characterization for more general A_1.

We will use $\|\cdot\|_m$ to denote any one of the equivalent norms on real Euclidean m-space R^m.

LEMMA 1. *If* $\{x_1, \ldots, x_k\}$ *is a linearly independent set of k vectors, each in* R^n, *then there exists* $\varepsilon > 0$ *and* $M > 0$ *such that if* $y_1, \ldots, y_k$ *are elements of* R^n *satisfying*

$$\|x_i - y_i\|_n < \varepsilon, \qquad i = 1, \ldots, k,$$

then

$$\left\{ \|\sigma\|_k : \sigma \in R^k \text{ and } \left\| \sum_{i=1}^{k} \sigma_i y_i \right\|_n \leqslant 1 \right\}$$

is bounded above by M.

Proof. By contradiction. Suppose that there exists a sequence of positive real numbers $\{\varepsilon_i\}$ which converges to zero, a sequence of subsets of R^n,

$$\{\{y_1^{(j)}, \ldots, y_k^{(j)}\}\}, \qquad j = 1, 2, \ldots,$$

and a sequence $\{\sigma^{(j)}\}$ in R^k such that

(a) $\|x_i - y_i^{(j)}\|_n < \varepsilon_j$, $i = 1, \ldots, k$,

(b) $\|\sigma^{(j)}\|_k \to \infty$ as $j \to \infty$ and

(c) $\|\sum_{i=1}^k \sigma_i^{(j)} y_i^{(j)}\|_n \leqslant 1$ for $j = 1, 2, \ldots$.

Let $\mu^{(j)} = \sigma^{(j)} / \|\sigma^{(j)}\|_k$.

Select a subsequence of $\{\mu^{(j)}\}$, call it $\{\mu^{(j_p)}\}$, which converges to a norm-one limit $\mu \in R^k$. By (b) and (c),

$$\sum_{i=1}^{k} \mu_i^{(j_p)} y_i^{(j_p)}$$

must converge to zero and also, by (a), to $\sum_{i=1}^k \mu_i x_i$. Hence $\sum_{i=1}^k \mu_i x_i$ equals the zero vector of R^n, a contradiction of the linear independence of $\{x_1,..., x_k\}$. This establishes Lemma 1.

LEMMA 2. (Discretization). *If* H_1–H_6 *are satisfied and if* $\{B_j\}$ *is a sequence of finite subsets of* A_1 *such that* $\max_{t\in A_1} \min_{\tau \in B_j} |t-\tau| \to 0$ *as* $j \to \infty$, *then* $\{h^{(j)}\}$ *converges uniformly to* h^* *on* $A_1 \cup A_2$, *where* $h^{(j)}$ *is the solution of* R_{B_j} *and* h^* *is the solution of* R_{A_1}.

Proof. Since h^* satisfies $v(t) \leqslant h^*(t) \leqslant u(t)$ for all t in B_j, and $h^{(j)}$ is the solution of R_{B_j}, we have

$$\int_{A_2} w(f-h^{(j)})^2 \leqslant \int_{A_2} w(f-h^*)^2 \qquad \text{for all } j. \tag{2.1}$$

Since all norms are equivalent on the finite dimensional space H, then $\{\max_{1\leqslant i\leqslant n} |a_i^{(j)}| : j=1, 2,...\}$ is bounded, where $h^{(j)} = \sum_{i=1}^n a_i^{(j)} h_i$. Hence $\{h^{(j)}\}$ is uniformly bounded on $A_1 \cup A_2$. Let $\{h^{(j_k)}\}$ be a subsequence which is uniformly convergent on $A_1 \cup A_2$, say to $\hat{h}$. It is easily seen that $v(t) \leqslant \hat{h}(t) \leqslant u(t)$ for all t in A_1. Also,

$$\int_{A_2} w(f-\hat{h})^2 = \lim_{k\to\infty} \int_{A_2} w(f-h^{(j_k)})^2 \leqslant \int_{A_2} w(f-h^*)^2$$

from (2.1). Hence the uniqueness of the solution of R_{A_1} implies that $\hat{h} = h^*$. Since $\{h^{(j)}\}$ is bounded and every convergent subsequence has h^* as limit, the entire sequence $\{h^{(j)}\}$ converges to h^*.

We now state some definitions and results on convex programming from Rockafellar [12].

Given a nonempty subset S of R^n, the *convex cone generated by* S is the set of all linear combinations of the form $\sum_{i=1}^p a_i s_i$, where $p > 0$, $a_i \geqslant 0$ and s_i in S for $i = 1,..., p$.

LEMMA 3. *Let* $\{S_i : i \in I\}$ *be a collection of nonempty convex sets in* R^n *and let* K *be the convex cone generated by* $\bigcup_{i\in I} S_i$. *Then every vector of* K *can be expressed as a linear combination with nonnegative coefficients of* n *or fewer linearly independent vectors, each belonging to a different* S_i.

Proof. [12, p. 156].

Recall that C is a *convex* subset of R^n if x, y in C and $0 < \theta < 1$ imply $\theta x + (1-\theta)y$ in C. A real function h is *convex* on a convex set $C \subseteq R^n$ if x, y in C and $0 < \theta < 1$ imply $h(\theta x + (1-\theta)y) \leqslant \theta h(x) + (1-\theta)\, h(y)$. If h differs from a linear function by a constant, it is *affine*.

An *(ordinary) convex program* (P) is a problem:

$$\begin{aligned} &\underset{x \in C}{\text{minimize}}\, f_0(x) \\ &\text{subject to}\, f_i(x) \leqslant 0, \qquad i = 1,\dots, r \\ &\qquad\qquad f_i(x) = 0, \qquad i = r+1,\dots, m, \end{aligned} \tag{P}$$

where C is a nonempty convex subset of R^n, $0 \leqslant r \leqslant m$, $f_1,\dots,f_r$ are convex functions on C and $f_{r+1},\dots,f_m$ are affine functions on C.

We say x in C is a *feasible solution* for (P) if x satisfies the m constraints of (P). The *optimal value* in (P) is $\inf(f_0(x)\colon x$ is a feasible solution for $(P)\}$. Of course, a feasible solution y is a *solution* of (P) if $f_0(y)$ equals this inf. The vector $\lambda = (\lambda_1,\dots, \lambda_m)$ is a *Kuhn–Tucker vector* for (P) if $\lambda_i \geqslant 0$ for $i = 1,\dots, r$ and if the infimum of $f_0 + \lambda_1 f_1 + \cdots + \lambda_m f_m$ over C is finite and equal to the optimal value in (P). The *Lagrangian* for (P) is the function $L(\lambda, x) = f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)$ defined for x in C and for $\lambda_i \geqslant 0$, $i = 1,\dots, r$. We say $(\bar\lambda, \bar x)$ is a *saddle point* for L if $L(\lambda, \bar x) \leqslant L(\bar\lambda, \bar x) \leqslant L(\bar\lambda, x)$ for all $(\lambda, \bar x)$ and $(\bar\lambda, x)$ in the domain of L.

LEMMA 4. *Let (P) be an ordinary convex program with $f_1,\dots,f_m$ affine and $C = R^n$. If the optimal value in (P) is not $-\infty$ and if (P) has a feasible solution, then a Kuhn–Tucker vector exists for (P).*

Proof. [12, p. 279].

In the next lemma the gradient vector is $\nabla h = (\partial h/\partial x_1,\dots, \partial h/\partial x_n)$.

LEMMA 5. *Let (P) be an ordinary convex program for which each f_i is differentiable. Let $\bar\lambda$ and $\bar x$ be vectors in R^m and R^n, respectively. In order that $\bar\lambda$ be a Kuhn–Tucker vector for (P) and $\bar x$ be an optimal solution of (P), it is necessary and sufficient that $(\bar\lambda, \bar x)$ be a saddle point for the Lagrangian of (P). Moreover this condition holds if and only if $\bar x$ and the components $\lambda_1,\dots, \lambda_m$ of $\bar\lambda$ satisfy*:

(a) $\lambda_i \geqslant 0$, $f_i(\bar x) \leqslant 0$, *and* $\lambda_i f_i(\bar x) = 0$ *for* $i = 1,\dots, r$,

(b) $f_i(\bar x) = 0$ *for* $i = r+1,\dots, m$,

(c) $\nabla(f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x))|_{x=\bar x} = 0$.

Proof. [12, pp. 280, 281].

We now prove our main result.

THEOREM 2. (Characterization). *Let $A_1, A_2, f, v, u, h_1, \ldots, h_n$ and w satisfy* H_1–H_7. *Then $h^* \in H$ is the solution of R_{A_1} if and only if*

(a) $v(t) \leqslant h^*(t) \leqslant u(t)$ *for all* $t \in A_1$ *and*

(b) *there exists a non-negative integer k, with $k \leqslant n$, distinct points $t_1, \ldots, t_k \in A_1$ and real constants $\sigma_1, \ldots, \sigma_k$ such that if* $1 \leqslant i \leqslant k$

(i) *either* $h^*(t_i) = u(t_i)$ *or* $h^*(t_i) = v(t_i)$,

(ii) *sign of* $\sigma_i = +1$ *if* $h^*(t_i) = u(t_i)$

$= -1$ *if* $h^*(t_i) = v(t_i)$,

(iii) *for* $j = 1, \ldots, n$,

$$\int_{A_2} w(t)(f(t) - h^*(t))\, h_j(t)\, dt = \sum_{i=1}^{k} \sigma_i h_j(t_i) \tag{2.2}$$

where the summation is interpreted as zero for $k = 0$.

Furthermore, (b) *is satisfied with* $k = 0$ *if and only if* h^* *is the best unconstrained least-squares approximation to* f *on* A_2.

Proof. The last statement is immediate, since for $k = 0$, (2.2) reduces to

$$\int_{A_2} w(f - h^*)\, h_j = 0, \qquad j = 1, \ldots, n,$$

which are the normal equations for unconstrained least-squares approximation and are well known to be necessary and sufficient for h^* to be a best approximation.

We now show that h^* in H is the solution of $R_{A_1} \Leftrightarrow h^*$ satisfies (a) and (b).

($\Rightarrow$) *Case* 1. A_1 has a finite number of points. Let $A_1 = \{t_1, \ldots, t_m\}$. Then R_{A_1} can be expressed equivalently as the convex programming problem

$$\underset{a}{\text{minimize}}\; Q(a) = \int_{A_2} w \left(f - \sum_{j=1}^{n} a_j h_j\right)^2 \text{ subject to}$$

$$c_i^-(a) \equiv v(t_i) - \sum_{j=1}^{n} a_j h_j(t_i) \leqslant 0 \text{ and}$$

$$c_i^+(a) \equiv \sum_{j=1}^{n} a_j h_j(t_i) - u(t_i) \leqslant 0, \text{ for } i = 1, \ldots, m.$$

We now apply Lemmas 4 and 5, which assure us that there exists a Kuhn–Tucker vector, $(\lambda_1^-, \ldots, \lambda_m^-, \lambda_1^+, \ldots, \lambda_m^+)$, of non-negative real numbers for R_{A_1} and that necessary and sufficient conditions for

$h^* = \sum_{j=1}^{n} a_j^* h_j$ to be a solution of R_{A_1} are

$$\lambda_i^- c_i^-(a^*) = 0 \text{ and } \lambda_i^+ c_i^+(a^*) = 0 \text{ for } i = 1,..., m$$

and

$$\nabla \left[Q(a) + \sum_{i=1}^{m} (\lambda_i^- c_i^-(a) + \lambda_i^+ c_i^+(a)) \right] \Bigg|_{a=a^*} = 0. \tag{2.3}$$

Now $\lambda_i^\pm \neq 0$ only if $c_i^\pm(a^*) = 0$; hence, by H_3, $\lambda_i^- \lambda_i^+ = 0$ for $i = 1,..., m$. Let

$$\begin{aligned} \sigma_i &= -0.5\lambda_i^- && \text{if } \lambda_i^- \neq 0 \\ &= 0.5\lambda_i^+ && \text{if } \lambda_i^+ \neq 0 \\ &= 0 && \text{if } \lambda_i^- = 0 \text{ and } \lambda_i^+ = 0. \end{aligned}$$

Equation (2.3) can be written as

$$\left[\frac{\partial}{\partial a_j} Q(a) \right] \Bigg|_{a=a^*} = - \left[\frac{\partial}{\partial a_j} \left(\sum_{i=1}^{m} \lambda_i^- c_i^-(a) + \lambda_i^+ c_i^+(a) \right) \right] \Bigg|_{a=a^*},$$

for $j = 1,..., n$. Or

$$-2 \int_{A_2} w(t)(f(t) - h^*(t))\, h_j(t)\, dt = \sum_{i=1}^{m} (\lambda_i^- h_j(t_i) - \lambda_i^+ h_j(t_i)),$$

for $j = 1,..., n$. Or, finally,

$$\int_{A_2} w(t)(f(t) - h^*(t))\, h_j(t)\, dt = \sum_{i=1}^{m} \sigma_i h_j(t_i),$$

for $j = 1,..., n$. Hence, if we pick out just the σ_i's that are non-zero and renumber them and their associated t_i's, we have established that h^* satisfies conditions (a) and (b).

We now prove $k \leqslant n$. Consider the convex cone generated by $\{\{v_i\}: i = 1,..., k\}$, where

$$v_i = (\sigma_i h_1(t_i),..., \sigma_i h_n(t_i)), \qquad \text{for } i = 1,..., k.$$

By Lemma 3,

$$x \equiv \left(\int_{A_2} w(t)(f(t) - h^*(t))\, h_1(t)\, dt,..., \int_{A_2} w(t)(f(t) - h^*(t))\, h_n(t)\, dt \right)$$

can be generated by a positive combination of no more than n of the v_i's. Hence, even if we started with more than n of the v_i's, all but at most n of these could be deleted. This concludes the proof of ($\Rightarrow$) for a finite A_1.

Case 2. A_1 has an infinite number of points. Since A_1 is bounded by H_1, we can select a sequence of finite sets $\{B_p\}$ such that

(1) $B_p \subseteq A_1$ for all p,

(2) $\max_{t \in A_1} \min_{s \in B_p} |t-s| \to 0$ as $p \to \infty$.

We apply Lemma 2 and see that the sequence of solutions of $\{R_{B_p}\}$, $\{h^{(p)}\}$, must converge to h^*, the solution of R_{A_1}. Note that h^* must satisfy (a) by definition; we will now show that it satisfies (b).

We will be selecting a number of subsequences. At each stage, we will use the subscript p_r for each of these to avoid cumbersome notation. For each p, $h^{(p)}$ satisfies (b) and $k=k_p$ is less than or equal to n. Select a subsequence of $\{R_{B_p}\}$ for which each k_p is equal to a common value, call it k. If $k=0$, then $\int_{A_2} w(f-h^*)\, h_j = \lim_{r\to\infty} \int w(f-h^{(p_r)})\, h_j = 0$ for $j=1,\dots,n$ and h^* satisfies (b). When $k \geqslant 1$ we index the $t_i^{(p_r)}$ (from condition (b) for $R_{B_{p_r}}$) so that

$$t_1^{(p_r)} < \cdots < t_k^{(p_r)}.$$

Select a subsequence for which $\{t_i^{(p_r)}\}$ converges for each $i=1,\dots,k$. This is possible by H_1. Some of these subsequences may coalesce. Let $t_1^*,\dots,t_q^*$ be the limit points of the k sequences, with $t_1^* < \cdots < t_q^*$ and $1 \leqslant q \leqslant k \leqslant n$.

Define the vectors

$$y_j^{(p_r)} = (h_1(t_j^{(p_r)}),\dots, h_n(t_j^{(p_r)})) \qquad \text{for } j=1,\dots,k,$$

and

$$x_i = (h_1(t_i^*),\dots, h_n(t_i^*)) \qquad \text{for } i=1,\dots,q.$$

Let

$$L_i = \{j \mid y_j^{(p_r)} \to x_i \text{ as } r \to \infty\}.$$

Condition (b)(ii) on the signs of the $\sigma_i^{(p_r)}$'s implies that if $j_1, j_2 \in L_i$, then sign $\sigma_{j_1}^{(p_r)} =$ sign $\sigma_{j_2}^{(p_r)}$ for all but a finite number of r's. Now

$$\sum_{j \in L_i} \sigma_j^{(p_r)}\, y_j^{(p_r)} = c_i^{(p_r)}\, z_i^{(p_r)},$$

where

$$c_i^{(p_r)} = \sum_{j \in L_i} \sigma_j^{(p_r)} \qquad \text{and} \qquad z_i^{(p_r)} = \frac{\sum_{j \in L_i} \sigma_j^{(p_r)}\, y_j^{(p_r)}}{\sum_{j \in L_i} \sigma_j^{(p_r)}}.$$

Notice

$$\|z_i^{(p_r)} - x_i\|_n = \left\| \sum_{j \in L_i} \left[\frac{\sigma_j^{(p_r)}}{\sum_{j \in L_i} \sigma_j^{(p_r)}} (y_j^{(p_r)} - x_i) \right] \right\|_n$$

$$\leqslant \sum_{j \in L_i} \| y_j^{(p_r)} - x_i \|_n \to 0 \qquad \text{as } r \to \infty.$$

Note that $x_1, \ldots, x_q$ are linearly independent by H_7.

The sequence of vectors, for $r = 1, 2, \ldots,$

$$\left\{ \left(\int_{A_2} w(t))f(t) - h^{(p_r)}(t))\, h_1(t)\, dt, \ldots, \int_{A_2} w(t)(f(t) - h^{(p_r)}(t))\, h_n(t)\, dt \right) \right\}$$

is convergent; hence, bounded. We now apply Lemma 1 to conclude that the sequence of vectors

$$\{c^{(p_r)}\} = \{(c_1^{(p_r)}, \ldots, c_q^{(p_r)})\}$$

must be bounded. Select a subsequence of $\{R_{B_{p_r}}\}$ for which $\{c^{(p_r)}\}$ is convergent to $c^* \in R^q$.

Let $t_{i_1}^*, \ldots, t_{i_q}^*$, be the t_i^*'s for which $c_i^* \neq 0$. We now have q' points in A_1 and q' real numbers for which h^* satisfies condition (b). This concludes the proof that the solution of R_{A_1} must satisfy (a) and (b).

$(\Leftarrow)$ If h^* satisfies (a) and (b), let $B = \{t_1, \ldots, t_k\}$ as in condition (b). By using Lemma 5 and reversing the analysis in Case 1 above, h^* must be the solution of R_B. If h in H satisfies $v(t) \leqslant h(t) \leqslant u(t)$ for all t in A_1, then h satisfies the constraints of R_B and

$$\int_{A_2} w(f - h^*)^2 \leqslant \int_{A_2} w(f - h)^2.$$

Since $v(t) \leqslant h^*(t) \leqslant u(t)$ for all t in A_1 by condition (a), h^* is the solution of R_{A_1}. This completes the proof of Theorem 2.

We note that if A_1 is a finite set, then the hypothesis H_7 is not needed for Theorem 2.

The characterization theorem can be used to check whether an approximation h^* which satisfies $v \leqslant h^* \leqslant u$ on A_1 is the solution of R_{A_1} as follows: Find $k \leqslant n$ points $t_1, \ldots, t_k$, where $h^* = u$ or $h^* = v$ and solve the linear equations (2.2) for $\sigma_1, \ldots, \sigma_k$ (it can be shown that a unique solution exists). If the signs of the resulting σ_i's are as prescribed in condition (b)(ii) of Theorem 2 then h^* is the solution of R_{A_1}.

We now present several corollaries which follow immediately from the characterization theorem.

In Corollary 1 we use the notation, for $B \subseteq A_1$,

$$E^2(B) = \int_{A_2} w(f - h_B)^2,$$

where h_B is the solution of the least-squares restricted range problem R_B.

COROLLARY 1. *If* H_1–H_7 *are satisfied and* h^* *is the solution of the least-squares restricted range problem* R_{A_1}, *then there exists a subset* T^* *of* A_1 *with at most n elements and such that the solution of* R_{T^*} *is* h^*. *Furthermore*

$$E^2(A_1) = E^2(T^*) = \max_{T \in \tau} E^2(T),$$

where τ *is the collection of all subsets of* A_1 *with at most n elements.*

Proof. Let $T^* = \{t_1, ..., t_k\}$ from Theorem 2, condition (b) for the problem R_{A_1}. Then h^* is the solution of R_{T^*} by Theorem 2 applied to R_{T^*}. This also shows $E^2(A_1) = E^2(T^*)$. Now if $T \subseteq A_1$, then any h in H which satisfies the constraints of R_{A_1} must satisfy the constraints of R_T.

Hence $E^2(A_1) \geqslant E^2(T)$.

COROLLARY 2. *If* H_1–H_7 *are satisfied and* h^* *is the solution of the least-squares restricted range problem* R_{A_1}, *then there exist a nonnegative integer* $k \leqslant n$, *a subset* $T = \{t_1, ..., t_k\}$ *of* A_1 *with* $t_1 < \cdots < t_k$ *and signs* $\varepsilon_1, ..., \varepsilon_k$, *each in* $\{-1, 1\}$, *such that* h^* *is the solution of the least-squares problem with interpolatory constraints*:

$$\begin{aligned} &\underset{h \in H}{\text{minimize}} \int_{A_2} w(f - h)^2 \\ &\textit{subject to } h(t_i) = u(t_i) \quad \text{if } \varepsilon_i = +1 \\ &\qquad\qquad\quad\;\; = v(t_i) \quad \text{if } \varepsilon_i = -1, i = 1, ..., k. \end{aligned} \tag{2.4}$$

Proof. Let $T = \{t_1, ..., t_k\}$ and $\sigma_1, ..., \sigma_k$ be as in Theorem 2 for the problem R_{A_1}; we may assume $t_1 < \cdots < t_k$. For $i = 1, ..., k$ set $\varepsilon_i =$ sign of σ_i. Since

$$\begin{aligned} h^*(t_i) &= u(t_i) \qquad \text{if} \quad \varepsilon_i = +1 \\ &= v(t_i) \qquad \text{if} \quad \varepsilon_i = -1 \end{aligned}$$

then h^* satisfies the constraints of problem (2.4). By the proof of Corollary

1, h^* is the solution of R_T. If h in H satisfies the constraints of problem (2.4), then h satisfies the constraints of R_T and so

$$\int_{A_2} w(f-h^*)^2 \leqslant \int_{A_2} w(f-h)^2.$$

Hence h^* is the solution of problem (2.4).

Theorem 2 gives a characterization condition for least-squares positive approximation and for least-squares one-sided approximation.

COROLLARY 3. *If* H_1, H_2, H_4, H_5 *and* H_7 *are satisfied, then* h^* *in* H *is the solution of the least-squares positive approximation problem*:

$$\underset{h\in H}{\text{minimize}} \int_{A_2} w(f-h)^2 \text{ subject to } h(t) \geqslant 0 \text{ for all } t \text{ in } A_1$$

if and only if

(a) $h^*(t) \geqslant 0$ *for all* t *in* A_1,

(b) *there exist a nonnegative integer* k, *with* $k \leqslant n$, *distinct points* $t_1,\dots, t_k$ *in* A_1, *and real constants* $\sigma_1,\dots, \sigma_k$ *such that*

(i) $h^*(t_i) = 0,\ i = 1,\dots, k,$

(ii) $\sigma_i < 0,\ i = 1,\dots, k,$

(iii) $\int_{A_2} w(f-h^*)\, h_j = \sum_{i=1}^{k} \sigma_i h_j(t_i), j = 1,\dots, n,$

where the summation is interpreted as zero for $k = 0$.

Proof. Let $v(t) \equiv 0$, $u(t) \equiv c > 0$. If c is chosen large enough, the upper constraint is inactive and the corollary follows from Theorem 2.

COROLLARY 4. *If* H_1, H_2, H_4, H_5 *and* H_7 *are satisfied and if there exists* h *in* H *with* $h(t) \leqslant f(t)$ *for all* t *in* A_1, *then* h^* *in* H *is the solution of the problem of least-squares one-sided approximation* (*from below*):

$$\underset{h\in H}{\text{minimize}} \int_{A_2} w(f-h)^2 \text{ subject to } h(t) \leqslant f(t) \text{ for all } t \text{ in } A_1$$

if and only if

(a) $h^*(t) \leqslant f(t)$ *for all* t *in* A_1,

(b) *there exist a nonnegative integer* k, *with* $k \leqslant n$, *distinct points* $t_1,\dots, t_k$ *in* A_1, *and real constants* $\sigma_1,\dots, \sigma_k$ *such that*

(i) $h^*(t_i) = f(t_i),\ i = 1,\dots, k$

(ii) $\sigma_i > 0,\ i = 1,\dots, k$

(iii) $\int_{A_2} w(f-h^*)\, h_j = \sum_{i=1}^{k} \sigma_i h_j(t_i), j = 1,\dots, n$

where the summation is interpreted as zero for $k = 0$.

Proof. Let $u(t)=f(t), v(t)\equiv C<0$. If $|C|$ is chosen large enough, the lower constraint is inactive and the corollary follows from Theorem 2.

Of course, the analogs of Corollaries 1 and 2 hold for positive approximation and for one-sided approximation.

We remark that the above analysis extends with very little change to give a characterization theorem for L_{2p} ($2p$ an even integer) approximation with restricted range:

$$\underset{h\in H}{\text{minimize}} \int_{A_2} w(f-h)^{2p} \text{ subject to } v(t)\leqslant h(t)\leqslant u(t) \text{ for all } t \text{ in } A_1.$$

Existence and uniqueness of a solution again follows using standard arguments. The characterization result is this:

THEOREM 3. (Characterization). *If* H_1–H_7 *are satisfied and if* $2p$ *is an even positive integer, then* h^* *in* H *is the solution of the* L_{2p} *approximation problem with restricted range if and only if*

(a) $v(t)\leqslant h^*(t)\leqslant u(t)$ *for all* t *in* A_1,

(b) *there exist a nonnegative integer* k, *with* $k\leqslant n$, *distinct points* $t_1,\ldots,t_k$ *in* A_1, *and real constants* $\sigma_1,\ldots,\sigma_k$ *such that if* $1\leqslant i\leqslant k$

(i) *either* $h^*(t_i)=u(t_i)$ *or* $h^*(t_i)=v(t_i)$ *for* $i=1,\ldots,k$,

(ii) *sign of* $\sigma_i=+1$ if $h^*(t_i)=u(t_i)$
$=-1$ if $h^*(t_i)=v(t_i)$,

(iii) *for* $j=1,\ldots,n$

$$\int_{A_2} w(f-h^*)^{2p-1} h_j=\sum_{i=1}^{k}\sigma_i h_j(t_i),$$

where the summation is interpreted as zero for $k=0$. *Furthermore* (b) *is satisfied with* $k=0$ *if and only if* h^* *in* H *is the best unconstrained* L_{2p} *approximation to* f *on* A_2.

Of course, the analogs of Corollaries 1–4 hold for L_{2p} approximation.

REFERENCES

1. F. DEUTSCH, J. H. MCCABE, AND G. M. PHILLIPS, Some algorithms for computing best approximation from convex cones, *SIAM J. Numer. Anal.* **12** (1975), 390–403.
2. R. DUFFIN AND L. KARLOVITZ, Formulation of linear programs in analysis. I: Approximation theory, *SIAM J. Appl. Math.* **16** (1968), 662–675.

3. R. J. EVANS AND A. CANTONI, A note on least squares approximation with restricted range constraints, *J. Austral. Math. Soc.* *B* **21** (1979), 230–242.
4. K. R. GEHNER, Necessary and sufficient optimality conditions for the Fritz John problem with linear equality constraints, *SIAM J. Control* **12** (1974), 140–149.
5. K. R. GEHNER, Characterization theorems for constrained approximation problems via optimization theory, *J. Approx. Theory* **14** (1975), 51–76.
6. E. KIMCHI AND N. RICHTER-DYN, Restricted range approximation of *k*-convex functions in monotone norms, *SIAM J. Numer. Anal.* **15** (1978), 1030–1038.
7. K. M. LEVASSEUR, Best approximation with respect to two objectives, Doctoral thesis, University of Rhode Island, 1980.
8. J. T. LEWIS, Approximation with convex constraints, *SIAM Rev.* **15** (1973), 193–217.
9. J. T. LEWIS, R. MURPHY, AND D. W. TUFTS, Design of minimum noise digital filters subject to inequality constraints using quadratic programming, *IEEE Trans. Acoust. Speech Signal Process.* (1976), 434–436.
10. J. W. MATHEWS, An algorithm for obtaining best approximations on sets of positive polynomials, *SIAM J. Math. Anal.* **9** (1972), 734–742.
11. J. R. RICE, Approximation with convex constraints, *J. SIAM* **11** (1963), 15–32.
12. R. T. ROCKAFELLAR, "Convex Analysis," Princeton Univ. Press, Princeton, N.J., 1970.
13. G. D. TAYLOR, Approximation by functions having restricted ranges, III, *J. Math. Anal. Appl.* **27** (1969), 241–248.