JOURNAL OF APPROXIMATION THEORY 14, 218–224 (1975)

Characterization and Uniqueness of Best Uniform Vector-Valued Linear Approximation

PAUL SCHWARTZ

Department of Mathematics, University of Texas, Austin, Texas 78712

Communicated by E. W. Cheney

1. INTRODUCTION

Given a real normed linear space X and a compact Hausdorff space Q, let $C(Q, X)$ denote the space of continuous functions from Q to X. We equip C with the uniform norm defined by $\| h \| = \max_{q \in Q} \| h(q) \|$, $h \in C$. Best linear approximation of functions in these normed function spaces has been systematically studied by Zuhovickiĭ, Krein, and Stečkin [6], and L. W. Johnson [3], [4]. Our aim in this note is to obtain characterization and uniqueness theorems analogous to those of the real-valued case (X = the real field $\mathbb{R}$) by a direct reduction to the known results of the real-valued situation, for the most part. I am much indebted to Professor E. W. Cheney for the suggestion that this should be possible.

2. CHARACTERIZATION THEOREMS

Let X^* denote the dual space, and $B(X^*)$ the closed unit ball of X^*. In what follows X^* carries the weak-star topology.

Given $f \in C(Q, X)$, define the *real-valued* function $\bar{f} \in C(Q \times B(X^*), \mathbb{R})$ by $\bar{f}(q, L) = L(f(q))$, $(q, L) \in Q \times B(X^*)$. (With the product topology, $Q \times B(X^*)$ is compact.) It is easy to verify that the uniform norm of $\bar{f}$, $\| \bar{f} \| = \max_{(q,L) \in Q \times B(X^*)} | \bar{f}(q, L) | = \| f \|$; and that the critical point set of $\bar{f}$, $\text{crit}(\bar{f}) = \{(q, L) \mid L(f(q)) = \pm \| f \| \}$. Note that $\bar{f}(q, -L) = -\bar{f}(q, L)$.

From this, we see at once that $f \mapsto \bar{f}$ is a norm-preserving linear map of $C(Q, X)$ into $C(Q \times B(X^*), \mathbb{R})$, so subspaces $V \subset C(Q, X)$ map onto subspaces $\bar{V} = \{\bar{v} \mid v \in V\} \subset C(Q \times B(X^*), \mathbb{R})$. Hence the problem of best linear approximation in $C(Q, X)$ is transferred into the corresponding real-valued problem in $C(Q \times B(X^*), \mathbb{R})$. Explicitly: given a subspace $V \subset C(Q, X)$ and $f \in C(Q, X)$, v_0 in V is a best approximation out of V to f if and only if $\bar{v}_0$ in $\bar{V}$ is a best approximation to $\bar{f}$.

In Theorems 1 and 2 to follow, we simply convert classic characterizations of real-valued best approximation into vector-valued ones.

THEOREM 1 (Kolmogorov-type characterization). *Let V be a subspace of $C(Q, X)$, $f \in C$. A necessary and sufficient condition that v_0 (in V) be a best approximation to f is the following*:

(K) *for each $v \in V$ there exists $(q, L) \in Q \times B(X^*)$ such that $L(f(q) - v_0(q)) = \|f - v_0\|$ and satisfying $L(v(q)) \leqslant 0$.*

Moreover, if for each v in V, $v \neq v_0$, there exists (q, L) with $L(f(q) - v_0(q)) = \|f - v_0\|$, and satisfying $L(v(q)) < 0$, v_0 is the unique best approximation.

Proof. In $C(Q \times B(X^*), \mathbb{R})$ $\bar{v}_0$ is a best approximation to $\bar{f}$ if and only if the well-known Kolmogorov condition holds: for each $\bar{v} \in \bar{V}$ there exists $(q, L) \in \text{crit}(\bar{f} - \bar{v}_0)$ satisfying $(\bar{f}(q, L) - \bar{v}_0(q, L)) \cdot \bar{v}(q, L) \leqslant 0$. Since $(q, -L)$ and (q, L) are simultaneously critical points of $\bar{f} - \bar{v}_0$, the asserted characterization follows.

We note that Theorem 1 at once implies the necessity (the sufficiency is clear) of the following characterization of best approximations involving only critical points (for $f \in C$, $\text{crit}(f) = \{q \in Q \mid \|f(q)\| = \|f\|\}$).

A necessary and sufficient condition that v_0 be best to f is that for each v in V there exists $q \in \text{crit}(f - v_0)$ such that $\|f - v_0\| \leqslant \|f(q) - v(q)\|$. With more effort, this characterization was obtained in the somewhat different setting of Johnson [3].

The characterization as formulated in Theorem 1 was obtained in a setting covering the case of X a smooth finite-dimensional normed space in Johnson [3], and by Zuhovickiĭ for X a Hilbert space (see [6]).

In practice, simpler characterizations may result from restricting the functionals in $B(X^*)$ to the set $\text{Ep}\, B(X^*)$ of extreme points of $B(X^*)$. I.e., condition (K) in Theorem 1 may be replaced by (K^e), where (K^e) is obtained from (K) by just replacing "$(q, L) \in Q \times B(X^*)$" by "$(q, L) \in Q \times \text{Ep}\, B(X^*)$." The following argument shows that this is possible. Given $q \in Q$, let $S(f(q) - v_0(q)) = \{L \in B(X^*) \mid L(f(q) - v_0(q)) = \|f(q) - v_0(q)\|\}$; this set is weak-star compact and an extremal subset of $B(X^*)$, so the extreme points of $S(f(q) - v_0(q))$ are precisely the extreme points of $B(X^*)$ in $S(f(q) - v_0(q))$. Now suppose L in $S(f(q) - v_0(q))$ satisfies $L(v(q)) \leqslant 0$, as in (K). The evaluation map $\varphi: L \mapsto L(v(q))$ is weak-star continuous and $a = \min_{L \in S(f(q) - v_0(q))} L(v(q)) \leqslant 0$. We now require the following result from extreme point theory (see Köthe [5, Section 25, p. 333]): (†) If T is a continuous linear map of one locally convex space E into another, then for any compact $C \subset E$, every extreme point of the image $T(C)$ is the image of some extreme point of C. a is an extreme point of $\varphi[S(f(q) - v_0(q))] \subset \mathbb{R}$, so by (†) there is an extreme point L' of $S(f(q) - v_0(q))$ satisfying $a = \varphi(L') = L'(v(q))$.

In the examples we now discuss, (K^e) will be used as the characterizing condition.

EXAMPLE 1. Uniform Approximation in $C(Q, \ell_\infty^n)$. The uniform norm is $\| h \| = \max_{q \in Q} \max_{1 \leqslant i \leqslant n} | h_i(q)|$, $h = (h_1, \ldots, h_n)$ in C. The dual $(\ell_\infty^n)^* = \ell_1^n$, so has extreme points $\{\pm e_i, i = 1, \ldots, n\}$ ($e_1, \ldots, e_n$ are the standard unit coordinate vectors of $\mathbb{R}^n$). So, $(q, L) \in Q \times \operatorname{Ep} B((\ell_\infty^n)^*)$ satisfies $L(h(q)) = \| h \|$ just in case $\| h(q)\|_\infty = \| h \|$, and $L \in \{\operatorname{sgn} h_j(q) \cdot e_j \mid | h_j(q)| = \| h(q)\|_\infty\}$.

With these observations recorded, Theorem 1 assumes the following concrete form.

Let V be a subspace of $C(Q, \ell_\infty^n)$, $f = (f_1, \ldots, f_n) \in C$. $v_0 = (v_{01}, \ldots, v_{0n})$ in V is a best approximation to f if and only if the following condition holds:

For each $v = (v_1, \ldots, v_n)$ in V there exist $q \in Q$ and $i = 1, \ldots, n$ such that $| f_i(q) - v_{0i}(q)| = \| f - v_0 \|$, satisfying $(f_i(q) - v_{0i}(q)) \cdot v_i(q) \leqslant 0$.

EXAMPLE 2 Uniform Approximation in $C(Q, \ell_1^n)$. The uniform norm is $\| h \| = \max_{q \in Q} (| h_1(q)| + \cdots + | h_n(q)|)$, h in C. The dual $(\ell_1^n)^* = \ell_\infty^n$, so has extreme points $\{\sum_{i=1}^n \sigma_i e_i : \sigma_i = \pm 1\}$. Hence $(q, L) \in Q \times \operatorname{Ep} B((\ell_1^n)^*)$ satisfies $L(h(q)) = \| h \|$ just in case $\| h(q)\|_1 = \| h \|$, and

$$L \in \left\{ \sum_{j : h_j(q) \neq 0} \operatorname{sgn} h_j(q) \cdot e_j + \sum_{i : h_i(q) = 0} \sigma_i \cdot e_i \,\middle|\, \sigma_i = \pm 1 \text{ for the } i \text{ such that } h_i(q) = 0 \right\}.$$

Theorem 1 here reads as follows.

Let V be a subspace of $C(Q, \ell_1^n)$, $f = (f_1, \ldots, f_n) \in C$. $v_0 = (v_{01}, \ldots, v_{0n})$ in V is a best approximation to f if and only if the following condition holds:

For each $v = (v_1, \ldots, v_n)$ in V there exists q such that $\| f(q) - v_0(q)\|_1 = \| f - v_0 \|$, satisfying

$$\sum_{j : f_j(q) - v_{0j}(q) \neq 0} \operatorname{sgn}(f_j(q) - v_{0j}(q)) \cdot v_j(q) \leqslant \sum_{i : f_i(q) - v_{0i}(q) = 0} | v_i(q)|.$$

THEOREM 2 ($0 \in$ *convex hull finite-dimensional characterization*). *Let V be an l-dimensional subspace of $C(Q, X)$. The following condition is necessary and sufficient that v_0 (in V) be a best approximation to f:*

There exist r elements $(q_1, L_1), \ldots, (q_r, L_r) \in Q \times B(X^*)$, $r \leqslant l + 1$, such that $L(f(q_i) - v_0(q_i)) = \| f - v_0 \|$, and barycentric coordinates $(\lambda_1, \ldots, \lambda_r)$ such that for all v in V, $\lambda_1 L_1(v(q_1)) + \cdots + \lambda_r L_r(v(q_r)) = 0$.

In addition, the elements (q_i, L_i) may be chosen in $Q \times \operatorname{Ep} B(X^*)$.

Proof. In $C(Q \times B(X^*), \mathbb{R})$, $\bar{V}$ is also l-dimensional and the classic $0 \in$ convex hull characterization that $\bar{v}_0$ in $\bar{V}$ be a best approximation to $\bar{f}$ reads: there are r $(\leqslant l+1)$ points $(q_1, L_1),..., (q_r, L_r) \in \text{crit}(\bar{f} - \bar{v}_0)$ and barycentric coordinates $(\lambda_1, ..., \lambda_r)$ satisfying (*)

$$\sum_{i=1}^{r} \lambda_i[\text{sgn}(\bar{f} - \bar{v}_0)(q_i, L_i)] \cdot \bar{v}(q_i, L_i) = 0 \qquad \text{for all} \quad \bar{v} \in \bar{V}.$$

Since (q_i, L_i) and $(q_i, -L_i)$ both belong to $\text{crit}(\bar{f} - \bar{v}_0)$, we choose the critical point (q_i, L_i) so that $(\bar{f} - \bar{v}_0)(q_i, L_i) = \|\bar{f} - \bar{v}_0\|$. (*) now translates to the condition given.

We now show how only extreme points need be used. Let V have basis $v_1, ..., v_l$, put $P_{(q,L)} = (Lv_1(q), ..., Lv(q))$, and let $U = \{P_{(q,L)} \mid (q, L) \in Q \times B(X^*)$ satisfies $L(f(q) - v_0(q)) = \|f - v_0\|\}$. U is a compact subset of $\mathbb{R}^n$, and what has been established so far simply shows that 0 is in the convex hull of U. By the finite-dimensional Krein–Milman theorem and Carathéodory's theorem, 0 is a convex linear combination of s $(\leqslant l+1)$ extreme points $P_{(q_1,L_1)}, ..., P_{(q_s,L_s)}$ of U. Fix $i = 1,..., s$ and consider the continuous linear map $L \mapsto (Lv_1(q_i), ..., Lv(q_i))$ from X^* to $\mathbb{R}^n$. Taking the compact $C = S(f(q_i) - v_0(q_i))$ in the result (†) cited above, it follows that there is an extreme point L_i' in $S(f(q_i) - v_0(q_i))$ such that $P_{(q_i,L_i)} = P_{(q_i,L_i')}$.

Adapting the terminology of Collatz [2] to the present vector-valued situation, call any compact subset $K \subset Q \times B(X^*)$ an H-set for the triple (V, v_0, f) if K satisfies the following two conditions:

(1) $L(f(q) - v_0(q)) = \|f - v_0\|$ for all $(q, L) \in K$,

(2) for each $v \in V$ there exists $(q, L) \in K$ satisfying $L(v(q)) \leqslant 0$.

With this terminology, Theorem 1 shows that (i) v_0 is a best approximation to f if and only if $\{(q, L) \mid L(f(q) - v_0(q)) = \|f - v_0\|\}$ is an H-set; (ii) If K is any H-set for (V, v_0, f), then v_0 is a best approximation to f.

After Theorem 2, just as in the real-valued situation, the Kolmogorov-type characterization of Theorem 1 for finite-dimensional V can be improved as follows. If $\dim V = l$, v_0 is a best approximation to f if and only if there is an H-set for (V, v_0, f) of cardinality $\leqslant l+1$. In addition, it is easy to verify that any H-set for (V, v_0, f) contains an H-set of cardinality $\leqslant l+1$ (consider the argument in the real-valued case that leads from Kolmogorov-type characterization to the $0 \in$ convex hull one in the finite-dimensional case). If $\{(q_1, L_1), ..., (q_r, L_r)\}$ is an H-set for (V, v_0, f) then v_0 is also a best approximation to f on $\{q_1, ..., q_r\}$ and $\|f - v_0\| = \min_{v \in V} \max_{1 \leqslant i \leqslant r} \|f(q_i) - v(q_i)\|$. So as for real-valued approximation, in principle a finite-dimensional linear vector-valued problem can be discretized.

3. The Uniqueness of Best Linear Approximation

Following Cheney and Wulbert [1], a subset K of Q is termed an α-set for subspace V if $K = \operatorname{crit}(f - v_0)$, where f (in C) has best approximation v_0 (in V). An α-set is thus the set of first coordinates of a certain H-set for V, and the H-sets built on α-sets are adequate to characterize any best approximation from V. Say that V satisfies condition (C) if no nonzero function in V vanishes identically on an α-set of V.

Theorem 3. *Condition* (C) *is necessary for best approximations out of V to be unique* (*when they exist at all*). *If X is strictly convex, condition* (C) *is also sufficient for unicity*; *but if X is not strictly convex, condition* (C) *is not sufficient.*

Proof. Assume best approximations are unique, but suppose K is an α-set on which some nonzero v_0 in V vanishes identically. Choose f in C such that $K = \operatorname{crit}(f)$ and f has 0 as best approximation. Define h in C by $h(q) = (\| v_0 \| - \| v_0(q) \|)(f(q)/\| f \|)$. We have $\| h \| = \| v_0 \|$ and $\{(q, L) \mid L(h(q)) = \| h \|\} = \{(q, L) \mid L(f(q)) = \| f \|\}$. From this, the Kolmogorov-type characterization of Theorem 1 shows that 0 is a best approximation to h. But $\| h(q) - v_0(q) \| \leqslant \| v_0 \|$, $q \in Q$, so also v_0 is a best approximation to h. Hence V satisfies condition (C).

Assume that X is strictly convex and condition (C) holds. Suppose f (not in V) has two best approximations v_1, v_0 in V. Clearly v_1 is a best approximation to $2f - v_1$, and as $\| 2f - v_1 - v_0 \| \leqslant 2 \| f - v_1 \|$, also v_0 is a best approximation to $2f - v_1$. $\| 2f(q) - v_1(q) - v_0(q) \| \leqslant \| f(q) - v_1(q) \| + \| f(q) - v_0(q) \| \leqslant 2 \| f - v_1 \|$, so for $q \in \operatorname{crit}(2f - v_1 - v_0)$ $\|(f(q) - v_1(q)) + (f(q) - v_0(q))\| = \| f(q) - v_1(q) \| + \| f(q) - v_0(q) \|$ and $\| f(q) - v_1(q) \| = \| f(q) - v_0(q) \|$. Since X is strictly convex, this implies $f(q) - v_0(q) = f(q) - v_1(q)$, so $v_1 - v_0$ vanishes on the α-set $\operatorname{crit}(2f - v_1 - v_0)$.

Finally assume X is not strictly convex. It is well known (and easy to see) that there is a one-dimensional subspace V and a point f not in V having more than one best approximation out of V. By taking the constant functions corresponding to V and f, we get a one-dimensional subspace of $C(Q, X)$ satisfying condition (C) but not admitting unique best approximations.

Stated as a complete characterization of subspaces admitting unique best approximations, Theorem 3 was first obtained by Cheney and Wulbert [1] for $X = \mathbb{R}$, and for a setting covering the case X strictly convex in Johnson [4]. Condition (C) deserves a little more examination. When V is finite-dimensional, it can readily be shown that condition (C) is equivalent to the zero and interpolation conditions on V given by Zuhovickiĭ, Krein, and Stečkin [6] as necessary and sufficient for V to admit unique best approximations (some relaxation of their additional restrictions on X is possible).

Further, if condition (C) is regarded as a uniqueness criterion, the class of α-sets is adequate to build H-sets characterizing any best approximation and to decide the uniqueness question. Another class of sets introduced below, related to α-sets by a minimal property, also performs both these functions.

By virtue of Zorn's lemma, two facts hold in general: given f having a best approximation v_0 in V, (1) any H-set for (V, v_0, f) contains a minimal H-set; (2) $\text{crit}(f - v_0)$ contains a compact set K minimal with respect to the property that v_0 is a best approximation to f on K. If V is l-dimensional, no transfinite argument is necessary and both a minimal H-set and K have cardinality $\leqslant l + 1$.

Given f having a best approximation v_0 in V, a compact set K having property (2) will be termed an η-set for V. η-sets are thus adequate as bases for H-sets characterizing best approximations, and further they can replace α-sets in condition (C)—i.e., condition (C) is equivalent to condition (C′): no nonzero member of V vanishes identically on an η-set. Clearly (C′) implies (C). If condition (C′) fails, choose an η-set K for which there is a nonzero v^* in V vanishing identically on K, which we can take with $\| v^* \| = 1$. Choose f in C having 0 as best approximation such that 0 is also a best approximation to f on K. Put $h(q) = (1 - \| v^*(q)\|) f(q)$. $\| h \| = \| f \|$ and $\text{crit}(h) = \text{crit}(f) \cap \{q \mid v^*(q) = 0\} \supset K$. Given v in V, there exists $(q, L) \in Q \times B(X^*)$ such that $L(f(q)) = \|f\|$, with $q \in K$, satisfying $L(v(q)) \leqslant 0$. $\| h(q) - v(q)\| \geqslant L(h(q) - v(q)) \geqslant L(h(q)) = L(f(q)) = \| h \|$, so h has 0 as best approximation and v^* vanishes identically on the α-set $\text{crit}(h)$. Hence (C) fails if (C′) does.

For finite-dimensional V, minimal H-sets can be identified independently of any approximation problem, so η-sets are thereby identified also. Such an identification runs as follows.

d elements $(q_1, L_1), \ldots, (q_d, L_d)$ of $Q \times B(X^*)$ form a minimal H-set for V if and only if the following conditions hold:

(1) $L_1, \ldots, L_d$ are norm-one functionals attaining their norm on the unit sphere of X;

(2) all the distinct functionals in pairs having the same first coordinate attain their norm at some one point of the unit sphere of X;

(3) there are unique positive barycentric coordinates $\lambda_1, \ldots, \lambda_d$ such that $\sum_{i=1}^{d} \lambda_i L_i(v(q_i)) = 0$ for all v in V.

(3) in turn can be given the following matrix formulation. Choosing a basis $v_1, \ldots, v_l$ for V, (3) holds just in case the $d \times l$ matrix $A = [L_i v_j(q_i)]_{i=1,\ldots,d;\, j=1,\ldots,l}$ has the following two properties:

(1) $\text{rank}\, A = d - 1$.

(2) There are $d - 1$ columns of A, say with indices $j_1 < \cdots < j_{d-1}$, such that for $i = 1, \ldots, d$ $\det A[1, \ldots, i-1, i+1, \ldots, d \mid j_1, \ldots, j_{d-1}] \neq 0$ and

all the $(-1)^i \det A[1,..., i-1, i+1,..., d \mid j_1,..., j_{d-1}]$ have the same sign. ($A[i_1,..., i_u \mid j_1,..., j_v]$ denotes the $u \times v$ submatrix of A formed from the intersections of the rows of A having indices $i_1 < \cdots < i_u$ with the columns of A having indices $j_1 < \cdots < j_v$.)

REFERENCES

1. E. W. CHENEY AND D. E. WULBERT, The existence and unicity of best approximations, *Math. Scand.* **24** (1969), 113–140.
2. L. COLLATZ, Inclusion theorems for the minimal distance in rational Tschebyscheff approximation with several variables, *in* "Approximation of Functions" (H. L. Garabedian, Ed.), 1965.
3. L. W. JOHNSON, Uniform approximation of vector-valued functions, *Numer. Math.* **13** (1969), 238–244.
4. L. W. JOHNSON, Unicity in the uniform approximation of vector-valued functions, *Bull. Austral. Math. Soc.* **3** (1970), 193–198.
5. G. KÖTHE, "Topological Vector Spaces I," New York, 1969.
6. S. I. ZUHOVICKIĬ AND S. B. STEČKIN, On the approximation of abstract functions, *Amer. Math. Soc. Transl.* **16** (1960), 401–406.