JOURNAL OF APPROXIMATION THEORY 29, 132–143 (1980)

Best Approximation with Prescribed Norm

G. A. WATSON

Department of Mathematics, University of Dundee, Dundee DD1 4HN, Scotland

Communicated by Oved Shisha

Received December 5, 1978

DEDICATED TO THE MEMORY OF P. TURÁN

1. INTRODUCTION

Let M be a real, normed linear space equipped with the norm $\|\cdot\|$, and let L be an n-dimensional linear subspace of M. If U denotes the unit sphere of L, defined by

$$U = \{f \in L: \|f\| = 1\},$$

then, for a given $m \in M$, an element $u^* \in U$ satisfying

$$\| u^* - m \| \leqslant \| u - m \|, \qquad \forall\, u \in U, \tag{1.1}$$

is a best approximation to m of unit norm. The problem of finding such a u^* is a particular example of a constrained best approximation problem. Such problems have received considerable attention in recent years, mainly with reference to the Chebyshev norm (see, for example, the review papers by Taylor [7] and Chalmers and Taylor[1]). For this norm, the present problem (with the assumption that L is a Haar subspace) has been studied by Ross and Belford [6], from the point of view of characterisation of local solutions (i.e., solutions u^* for which the inequality in (1.1) is satisfied for all $u \in U \cap N(u^*)$, with $N(u^*)$ a neighbourhood of u^*). The application of optimization (rather than more conventional approximation theoretic) techniques to general classes of nonlinear finite-dimensional inequality constrained best approximation problems is considered in [3, 8], and various conditions for a solution are derived. It is the object of this paper to apply similar techniques to the prescribed norm problem defined (without loss of generality) above.

If L is spanned by $l_1, l_2, ..., l_n$, then any element of L may be expressed in the form

$$l(\mathbf{a}) = \sum_{i=1}^{n} a_i l_i, \qquad \mathbf{a} \in R^n. \tag{1.2}$$

0021-9045/80/060132–12$02.00/0

The problem may then be restated as follows:

$$\begin{aligned}&\text{find } \mathbf{a} \in R^n \text{ to minimize } \| r(\mathbf{a})\|,\\ &\text{where } r(\mathbf{a}) = l(\mathbf{a}) - m\\ &\text{and } \| l(\mathbf{a})\| = 1.\end{aligned} \tag{1.3}$$

If there are only isolated points satisfying the constraint $\| l(\mathbf{a})\| = 1$ (for example, if $n = 1$, $M = R^1$), then each of these points gives, by definition, a local best approximation. We exclude this case therefore, and are thus able to define *feasible directional sequences*. Let $\{\mathbf{a}^{(k)}\} \to \mathbf{a}^*$ be a sequence of points in R^n such that if

$$\mathbf{a}^{(k)} - \mathbf{a}^* = \delta^{(k)}\mathbf{s}^{(k)}, \qquad \| \mathbf{s}^{(k)}\| = 1, \quad \delta^{(k)} > 0 \tag{1.4}$$

defines $\delta^{(k)}$ and $\mathbf{s}^{(k)}$ (where the norm is any norm on R^n), then $\mathbf{s}^{(k)}$ converges, to $\mathbf{s}$, say. Then $\{\mathbf{a}^{(k)}\}$ is a directional sequence. If $\| l(\mathbf{a}^{(k)})\| = 1$, then this sequence is feasible, and $\mathbf{s}$ is said to be a *feasible direction* at $\mathbf{a}^*$ with respect to the constraint of (1.3).

Now let M^{D} be the dual space of M, and $\| \cdot \|^{\mathrm{D}}$ the dual norm on M^{D}. Then M^{D} is the space of continuous linear functionals $v(m)$ defined on M, and it is convenient to write

$$v(m) = \langle m, v\rangle \tag{1.5}$$

thus expressing the linear functional as an inner product between the elements of M and those of M^{D}. An important role is played in what follows by the set of *subgradients* of $\| m \|$ at m, which is the set of elements $v \in M^{\mathrm{D}}$ which satisfy

$$\| m \| - \| n \| \leqslant \langle m - n, v\rangle \qquad \forall\, n \in M. \tag{1.6}$$

If this set is denoted by $V(m)$, then it is readily shown that we have

$$V(m) = \{v \in M^{\mathrm{D}}\colon \| m \| = \langle m, v\rangle, \| v \|^{\mathrm{D}} \leqslant 1\}. \tag{1.7}$$

The relationship between appropriate subgradients and the directional derivatives of the functions occurring in (1.3) is a crucial one, and the following result (which essentially generalizes Lemma 3 of [3]) is of fundamental importance. For convenience, the abbreviations $r^{(k)}$, l^*, etc., will be used to denote $r(\mathbf{a}^{(k)})$, $l(\mathbf{a}^*)$, etc., whenever they occur throughout this paper.

LEMMA 1. *Let* $\{\mathbf{a}^k\} \to \mathbf{a}^*$ *be a directional sequence defining* $\mathbf{s}$ *through* (1.4).

Then

$$\lim_{k\to\infty} \frac{\| r^{(k)} \| - \| r^* \|}{\delta^k} = \max_{v\in V(r^*)} \langle l(\mathbf{s}), v\rangle.$$

Proof. For all $v \in V(r^*)$ we have from (1.6)

$$\begin{aligned} \| r^{(k)}\| - \| r^* \| &\geqslant \langle r^{(k)} - r^*, v\rangle \\ &= \delta^{(k)}\langle l(\mathbf{s}^{(k)}), v\rangle. \end{aligned}$$

Also for $v^{(k)} \in V(r^{(k)})$,

$$\begin{aligned} \| r^{(k)}\| &= \langle r^{(k)}, v^{(k)}\rangle \\ &= \langle r(\mathbf{a}^* + \delta^{(k)}\mathbf{s}^{(k)}), v^{(k)}\rangle \\ &\leqslant \| r^* \| + \delta^{(k)}\langle l(\mathbf{s}^{(k)}), v^{(k)}\rangle. \end{aligned}$$

Thus

$$\langle l(\mathbf{s}^{(k)}), v^{(k)}\rangle \geqslant \frac{\| r^{(k)} \| - \| r^* \|}{\delta^{(k)}} \geqslant \langle l(\mathbf{s}^{(k)}), v\rangle \qquad \text{for all } v \in V(r^*). \quad (1.8)$$

By the weak * compactness of the unit ball in M^{D} (Alaoglu–Bourbaki theorem, e.g., Holmes [4]), there exists a sequence $\{\delta^{(k)}\} \to 0$ and $w \in M^{\mathrm{D}}$ such that

$$\langle m, v^{(k)}\rangle \to \langle m, w\rangle \qquad \text{as} \quad k \to \infty$$

for all $m \in M$. Further

$$0 \leqslant \| r^* \| - \langle r^*, v^{(k)}\rangle \leqslant O(\delta^{(k)})$$

and so $w \in V(r^*)$. Letting $k \to \infty$ along any such appropriate sequence, the result follows from the inequalities (1.8).

2. Conditions for a Best Approximation

Let $\mathscr{F}^*$ define the set of feasible directions for (1.3) at the (feasible) point $\mathbf{a}^*$, and define the sets

$$\begin{Bmatrix} F_P^* \\ F^* \\ F_N^* \end{Bmatrix} = \left\{ \mathbf{s} \in R^n : \| \mathbf{s} \| = 1,\ \max_{v\in V(l^*)} \langle l(\mathbf{s}), v\rangle \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \right\}.$$

Lemma 2. $F^* = \mathscr{F}^*$.

Proof. Let $\mathbf{s} \in \mathscr{F}^*$. Then there exists a feasible directional sequence $\{\mathbf{a}^{(k)}\} \to \mathbf{a}^*$. Thus by Lemma 1 for the special case when $m = 0$, $\mathbf{s} \in F^*$
Now let $\mathbf{s} \in F^*$. Then

$$\max_{v \in V(l^*)} \langle l(\mathbf{s}), v \rangle = 0. \tag{2.1}$$

Let $\mathbf{a}(\delta) = \mathbf{a}^* + \delta\mathbf{s}$, $\delta \geqslant 0$. Then $\{\mathbf{a}(\delta^{(k)})\}$ is a directional sequence, where $\delta^{(k)} \downarrow 0$ as $k \to \infty$. If $\| l(\mathbf{a}(\delta^{(k)}))\| = 1$ for k sufficiently large, then the directional sequence is feasible and the required result follows. Otherwise, let $\| l(\mathbf{a}(\delta^{(k)}))\| = \alpha^{(k)}$ and let $\mathbf{c}^{(k)} = \mathbf{a}(\delta^{(k)})/\alpha^{(k)}$, which is defined for all k sufficienly large. Then $\| l(\mathbf{c}^{(k)})\| = 1$, and it remains to show that $\{\mathbf{c}^{(k)}\} \to \mathbf{a}^*$ is a directional sequence. Now by Lemma 1, with m set equal to zero again,

$$\lim_{k \to \infty} \frac{\| l(\mathbf{a}(\delta^{(k)}))\| - \| l^* \|}{\delta^{(k)}} = 0$$

and so

$$\lim_{k \to \infty} \frac{| \alpha^{(k)} - 1 |}{\delta^{(k)}} = 0.$$

Also

$$\left\| \frac{\mathbf{c}^{(k)} - \mathbf{a}^*}{\delta^{(k)}} - \mathbf{s} \right\| = \left\| \frac{\mathbf{c}^{(k)} - \mathbf{a}(\delta^{(k)})}{\delta^{(k)}} \right\|$$

$$= \frac{| 1 - \alpha^{(k)} |}{\delta^{(k)}} \frac{\| \mathbf{a}(\delta^{(k)})\|}{\alpha^{(k)}}.$$

Thus $\mathbf{s} \in \mathscr{F}^*$.

LEMMA 3. *Let* $\mathbf{a}^*$ *give a best approximation. Then*

$$\max_{v \in V(r^*)} \langle l(\mathbf{s}), v \rangle \geqslant 0 \qquad \forall\, \mathbf{s} \in F^*.$$

Proof. Assume $\mathbf{a}^*$ gives a best approximation, and let $\mathbf{s} \in F^*$. Then $\mathbf{s} \in \mathscr{F}^*$ by Lemma 2, and so there exists a feasible directional sequence $\{\mathbf{a}^{(k)}\} \to \mathbf{a}^*$ with

$$\| r^{(k)} \| \geqslant \| r^* \|$$

for all k sufficiently large. Thus

$$\lim_{k \to \infty} \frac{\| r^{(k)} \| - \| r^* \|}{\delta^{(k)}} \geqslant 0$$

and the result follows from Lemma 1.

LEMMA 4. *If* $\mathbf{a}^*$ *is feasible, and if*

$$\max_{v \in V(r^*)} \langle l(\mathbf{s}), v \rangle > 0 \qquad \forall\, \mathbf{s} \in F^* \tag{2.2}$$

then $\mathbf{a}^*$ *gives an isolated local best approximation, in the sense that there exists a neighbourhood,* $N(l^*)$, *of* l^* *such that* $\| l^* - m \| < \| l(\mathbf{a}) - m \|$ *for all* $l(\mathbf{a}) \in U \cap N(l^*)$, $\mathbf{a} \neq \mathbf{a}^*$.

Proof. Let (2.2) be satisfied, let $\mathbf{a}^*$ be feasible, but suppose that $\mathbf{a}^*$ does not give an isolated local best aproximation. Then there exists a feasible sequence, and hence a feasible directional sequence $\{\mathbf{a}^{(k)}\} \to \mathbf{a}^*$ such that

$$\| r^{(k)} \| \leqslant \| r^* \|$$

for all k sufficiently large. Lemma 1 gives a contradiction and the result follows.

It is convenient at this point to introduce two separate cases of problem (1.3). Define the set G by

$$G = \{\mathbf{g} \in R^n \colon \| r(\mathbf{g})\| \leqslant \| r(\mathbf{a})\| \qquad \forall\, \mathbf{a} \in R^n\},$$

i.e., G is the set of vectors such that $l(\mathbf{g})$ is a best *unconstrained* approximation. The following characterization of G is required later (see, for example, [9] for a proof).

LEMMA 5. $\mathbf{g} \in G$ iff $\exists\, v \in V(r(\mathbf{g}))$ *such that* $\langle l_i, v \rangle = 0$, $i = 1, 2,\ldots, n$.

If $\| l(\mathbf{g})\| \geqslant 1$ for all $\mathbf{g} \in G$, then problem (1.3) is *precisely equivalent* to:

$$\begin{aligned} &\text{find } \mathbf{a} \in R^n \text{ to minimize } \| r(\mathbf{a})\| \\ &\text{subject to } \| l(\mathbf{a})\| \leqslant 1. \end{aligned} \tag{2.3}$$

With some modifications, the analysis of [3] is now directly applicable to this problem: since (2.3) is a convex programming problem, necessary and sufficient conditions for a solution may readily be obtained.

THEOREM 1. *Let* $\| l(\mathbf{g})\| \geqslant 1$ *for all* $\mathbf{g} \in G$. *Then* $\mathbf{a}^*$ *solves* (1.3) *iff* $\exists\, v \in V(r^*)$, $w \in V(l^*)$, $\lambda \geqslant 0$ *such that*

$$\langle l_i, v \rangle + \lambda \langle l_i, w \rangle = 0, \qquad i = 1, 2,\ldots, n. \tag{2.4}$$

Proof. Let $\mathbf{a}^*$ be a solution. If $\mathbf{a}^* \in G$ then the desired result holds with $\lambda = 0$ for some $v \in V(r^*)$ by Lemma 5. Thus we assume $\mathbf{a}^* \notin G$. We first show that

$$\max_{v \in V(r^*)} \langle l(\mathbf{s}), v \rangle \geqslant 0 \tag{2.5}$$

for all $\mathbf{s} \in F^* \cup F_N^*$. If $\mathbf{s} \in F^*$ this inequality is a consequence of Lemma 3 so let $\mathbf{s} \in F_N^*$. Then

$$\lim_{k\to\infty} \frac{\| l^{(k)} \| - \| l^* \|}{\delta^{(k)}} < 0$$

using Lemma 1, where $\{\mathbf{a}^{(k)}\} \to \mathbf{a}^*$ is a directional sequence, so that

$$\| l^{(k)} \| < \| l^* \| = 1$$

for k sufficiently large. Now if (2.5) is not satisfied, Lemma 1 gives that

$$\| r^{(k)} \| < \| r^* \|$$

for k sufficiently large. Thus the fact that $\mathbf{a}^*$ solves (2.3) is contradicted, and so (2.5) must hold as required.

Now define the closed, bounded, convex sets in R^n

$$B = \{\mathbf{b}: b_i = \langle -l_i, v\rangle, i = 1, 2,..., n, v \in V(r^*)\},$$
$$D = \{\mathbf{d}: d_i = \langle l_i, w\rangle, i = 1, 2,..., n, w \in V(l^*)\}$$

and let K be the convex cone in R^n generated by D, i.e.,

$$K = \{\mathbf{k}: \mathbf{k} = \alpha\mathbf{d}, \mathbf{d} \in D, \alpha \geqslant 0\}.$$

Now D does not contain the origin (for by Lemma 5 applied with $m = 0$ that would imply that $\mathbf{a}^*$ minimizes $\| l(\mathbf{a})\|$, a contradiction). Thus K is closed, and so by a standard separation result (for example, Lemma 6 of [3]) if $K \cap B = \varnothing$, there exists $\mathbf{s} \in R^n$ such that

$$\mathbf{s}^T\mathbf{k} \leqslant 0 \qquad \forall\, \mathbf{k} \in K,$$
$$\mathbf{s}^T\mathbf{b} > 0 \qquad \forall\, \mathbf{b} \in B.$$

Thus $\exists\, \mathbf{s} \in F^* \cup F_N^*$ for which (2.5) does not hold, which shows that $K \cap B \neq \varnothing$, and establishes the "only if" part of the theorem.

Now let the conditions (2.4) be satisfied at $\mathbf{a}^*$, and let $\mathbf{a} \in R^n$ be any other feasible point. Then for any $v \in V(r^*)$, $w \in V(l^*)$, $\lambda \geqslant 0$ satisfying (2.4) we have

$$\begin{aligned} \| r(\mathbf{a})\| - \| r^* \| &\geqslant \langle l(\mathbf{a}) - l^*, v\rangle \\ &= -\lambda\langle l(\mathbf{a}) - l^*, w\rangle \\ &= -\lambda\langle l(\mathbf{a}), w\rangle + \lambda \\ &= \lambda(1 - \langle l(\mathbf{a}), w\rangle) \geqslant 0. \end{aligned}$$

Remark. The proof of sufficiency of (2.4) does not require any assumption about the set G.

If $\| l(\mathbf{g})\| < 1$ for some $\mathbf{g} \in G$, then (1.3) is no longer equivalent to a convex programming problem: there may exist local best approximations which are not global, and in addition it is not always possible to close the gap between necessary conditions and sufficient conditions. In order to obtain necessary conditions analogous to those in Theorem 1, we require the following strengthened form of Lemma 3.

LEMMA 6. *Let* $\| l(\mathbf{g})\| < 1$ *for some* $\mathbf{g} \in G$, *and let* $\mathbf{a}^*$ *solve* (1.3) *with* $\mathbf{a}^* \notin G$. *Then*

$$\max_{v \in V(r^*)} \langle l(\mathbf{s}), v\rangle \geqslant 0 \qquad \forall\ \mathbf{s} \in F^* \cup F_P^* .$$

Proof. If $\mathbf{s} \in F^*$, the inequality follows from Lemma 3, so let $\mathbf{s} \in F_P^*$. Then using Lemma 1, if $\{\mathbf{a}^{(k)}\} \to \mathbf{a}^*$ is a directional sequence defining $\mathbf{s}$,

$$\| l^{(k)} \| > \| l^* \| = 1 \tag{2.6}$$

for k sufficiently large. Assume that

$$\max_{v \in V(r^*)} \langle l(\mathbf{s}), v\rangle < 0.$$

Then

$$\| r^{(k)} \| < \| r^* \| \tag{2.7}$$

for k sufficiently large, by Lemma 1. Also, by assumption, $\exists\ \mathbf{g} \in G$ such that

$$\| r(\mathbf{g})\| < \| r^* \| \tag{2.8}$$

with

$$\| l(\mathbf{g})\| < 1. \tag{2.9}$$

Thus from (2.6) and (2.9), $\exists\ \lambda^{(k)}$, $0 < \lambda^{(k)} < 1$, such that

$$\| l(\mathbf{a}(\lambda^{(k)}))\| = 1,$$

where $\mathbf{a}(\lambda^{(k)}) = \lambda^{(k)}\mathbf{a}^{(k)} + (1 - \lambda^{(k)})\, \mathbf{g}$. Further by (2.7) and (2.8)

$$\| r(\mathbf{a}(\lambda^{(k)}))\| < \| r^* \|,$$

and since $\mathbf{a}(\lambda^{(k)}) \to \mathbf{a}^*$ as $k \to \infty$, we contradict the fact that $\mathbf{a}^*$ gives a local best approximation, and the result follows.

THEOREM 2. *Let* $\| l(\mathbf{g})\| < 1$ *for some* $\mathbf{g} \in G$, *and let* $\mathbf{a}^*$ *solve* (1.3). *Then for* each $w \in V(l^*)$, $\exists\ v \in V(r^*)$, $\lambda \leqslant 0$ *such that*

$$\langle l_i , v\rangle + \lambda\langle l_i , w\rangle = 0, \qquad i = 1, 2,\dots, n. \tag{2.10}$$

Proof. Let $\mathbf{a}^*$ solve (1.3). Then if $\mathbf{a}^* \in G$, (2.10) is satisfied for some $v \in V(r^*)$ with $\lambda = 0$, by Lemma 5. Thus we consider the case when $\mathbf{a}^* \notin G$, and we will assume that the conditions (2.10) do not hold. Then $\exists\, w^* \in V(l^*)$ such that no $v \in V(r^*)$, $\lambda \leqslant 0$ satisfying (2.10) exist. Let K denote the set

$$K = \{\mathbf{k}: \mathbf{k} = \alpha\mathbf{d},\ d_i = \langle l_i, w^*\rangle,\ i = 1, 2,..., n,\ \alpha \geqslant 0\}.$$

Now $\mathbf{d} \neq 0$ (or else, by Lemma 5, applied with $m = 0$, we have the conclusion that $\mathbf{a}^*$ minimizes $\| l(\mathbf{a})\|$, a contradiction), and so K is a closed, half-line. Let B denote the closed, bounded convex set

$$B = \{\mathbf{b}: b_i = \langle l_i, v\rangle,\ i = 1, 2,..., n,\ v \in V(r^*)\}.$$

Then $K \cap B = \varnothing$ by assumption, and so $\exists\, \mathbf{s} \in R^n$ such that

$$\mathbf{s}^T\mathbf{k} \geqslant 0 \qquad \forall\, \mathbf{k} \in K,$$
$$\mathbf{s}^T\mathbf{b} < 0 \qquad \forall\, \mathbf{b} \in B.$$

This contradicts Lemma 6, and the result follows.

Theorems 1 and 2 taken together show that if $\mathbf{a}^*$ solves (1.3), $\exists\, v \in V(r^*)$, $w \in V(l^*)$ and a scalar λ such that

$$\langle l_i, v\rangle + \lambda\langle l_i, w\rangle = 0, \qquad i = 1, 2,..., n.$$

This particular form of necessary conditions gives a natural generalization of the differentiable case. Further, if $\lambda \geqslant 0$, then this condition is sufficient for $\mathbf{a}^*$ to solve (1.3). The conditions of Theorem 2 may in certain circumstances also be sufficient for a local best approximation. We have the following result.

THEOREM 3. *Let* $\mathbf{a}^*$ *be a feasible point such that*

(i) *for each* $w \in V(l^*)$, $\exists\, v \in V(r^*)$ *and* $\lambda \leqslant 0$ *such that*

$$\langle l_i, v\rangle + \lambda\langle l_i, w\rangle = 0, \qquad i = 1, 2,..., n \tag{2.11}$$

(ii) *there exists a neighbourhood* $N(\mathbf{a}^*)$ *of* $\mathbf{a}^*$ *such that*

$$V(l(\mathbf{a})) \subset V(l^*), \qquad \forall\, \mathbf{a} \in N(\mathbf{a}^*). \tag{2.12}$$

Then $\mathbf{a}^*$ *is a local solution of* (1.3).

Proof. Let $\{\mathbf{a}^{(k)}\}$ be any feasible sequence converging to $\mathbf{a}^*$, a point for which the given conditions (i) and (ii) are both satisfied. Then $\exists\, w^* \in V(l^*)$ such that

$$\langle l^{(k)}, w^*\rangle = \| l^{(k)} \| = 1$$

for k sufficiently large. Let $v^* \in V(r^*)$, λ^* and w^* satisfy (2.11). Then

$$\begin{aligned} \| r^{(k)} \| - \| r^* \| &\geqslant \langle l^{(k)} - l^*, v^* \rangle \\ &= -\lambda^* \langle l^{(k)} - l^*, w^* \rangle \\ &= 0. \end{aligned}$$

This theorem is particularly useful when (1.1) is set in a finite-dimensional space normed by a *polyhedral norm*. Consider the consistent set of linear inequalities

$$B\mathbf{u} \leqslant \mathbf{e},$$

where $\mathbf{u} \in R^t$, B is an $N \times t$ matrix, and $\mathbf{e} \in R^N$ is a vector, each component of which is 1. Then if

(i) $C = \{\mathbf{u}: B\mathbf{u} \leqslant \mathbf{e}\}$ is bounded and has a nonvoid interior,

(ii) $\mathbf{u} \in C$ if and only if $-\mathbf{u} \in C$,

the polyhedral norm on R^t specified by B is defined by

$$\| \mathbf{u} \| = \min\{\mu: B\mathbf{u} \leqslant \mu\mathbf{e}\}.$$

THEOREM 4. *Let $M = R^t$, normed by a polyhedral norm defined by the $N \times t$ matrix B. If condition* (i) *of Theorem* 3 *is satisfied at a feasible point* $\mathbf{a}^*$, *then* $\mathbf{a}^*$ *is a local solution of* (1.3).

Proof. It suffices to show that condition (ii) of Theorem 3 is automatically satisfied. Now at $\mathbf{a}^*$, let I^* be defined by

$$\begin{aligned} \langle b_i, r^* \rangle &= \| r^* \|, \qquad i \in I^*, \\ \langle b_i, r^* \rangle &< \| r^* \|, \qquad i \notin I^*, \end{aligned}$$

where $b_i \in R^t$ denotes the ith row of B. Then $V(r^*)$ is the convex hull of the set $\{b_i, i \in I^*\}$ (see, for example, [3] for a proof of this). Thus $\exists\, N(\mathbf{a}^*)$ for which (2.12) is satisfied and the result follows.

If $V(l^*)$ and $V(r^*)$ both contain *unique* elements v^* and w^*, it follows that $\| l(\mathbf{a})\|$ and $\| r(\mathbf{a})\|$ are differentiable at $\mathbf{a}^*$ (Rockafellar [5]). If, in addition, these functions are twice continuously differentiable in a neighbourhood of $\mathbf{a}^*$, then standard results from (differentiable) optimization theory are available to supplement those given earlier. In particular, if the norm is smooth, then such second-order conditions may be obtained in a natural manner. If it is known that $\| l(\mathbf{g})\| \geqslant 1$ for all $\mathbf{g} \in G$, such conditions are of course redundant; however, their derivation, and application, does not require such information. For completeness, we quote the relevant results as they apply in this particular case; see Fiacco and McCormick [2] for details.

Let $v^* \in V(r^*)$, $w^* \in V(l^*)$, $\lambda^* < 0$ satisfy

$$\langle l_i, v^* \rangle + \lambda^* < l_i, w^* \rangle = 0, \qquad i = 1, 2, ..., n, \tag{2.13}$$

and define

$$\mathscr{L}(\mathbf{a}) = \| r(\mathbf{a}) \| + \lambda^* \| l(\mathbf{a}) \|,$$
$$Z^* = \{\mathbf{s} \in R^n: \| \mathbf{s} \| = 1, \langle l(\mathbf{s}), v^* \rangle = \langle l(\mathbf{s}), w^* \rangle = 0\}.$$

THEOREM 5. *Let* $\| l(\mathbf{a}) \|$ *and* $\| r(\mathbf{a}) \|$ *be twice continuously differentiable in a neighbourhood of* $\mathbf{a}^*$, *a feasible point satisfying* (2.13). *Then*

(i) *if* $\mathbf{a}^*$ *solves* (1.3), *then*

$$\mathbf{s}^T \nabla^2 \mathscr{L}(\mathbf{a}^*) \, \mathbf{s} \geqslant 0 \qquad \forall \, \mathbf{s} \in Z^*,$$

(ii) if

$$\mathbf{s}^T \nabla^2 \mathscr{L}(\mathbf{a}^*) \, \mathbf{s} > 0 \qquad \forall \, \mathbf{s} \in Z^*,$$

then $\mathbf{a}^*$ is an isolated local solution of (1.3).

3. EXAMPLES

We conclude with some examples intended to illustrate the application of some of the results of the previous section, for the case where $M = C[-1, 1]$, normed with the L_∞ norm (see also [6]).

EXAMPLE 1.

$$n = 2, \qquad l_1 = 1, \qquad l_2 = x^2, \qquad m = \tfrac{1}{4}x^4.$$

Let $\mathbf{a}^* = (1, -1)^T$. Then

$$\| l^* \| = \| 1 - x^2 \| = 1$$

with $V(l^*) = \{\delta(0)\}$, where δ is the delta function defined for $f(x) \in M$ by

$$\langle f(x), \delta(\xi) \rangle = f(\xi).$$

Further

$$\| r^* \| = \| 1 - x^2 - \tfrac{1}{4} x^4 \| = 1$$

with $V(r^*) = \{\delta(0)\}$. Necessary conditions are therefore satisfied at $\mathbf{a}^*$ with $\lambda = -1$. Further, any perturbation of $\mathbf{a}^*$ will result in $\| l(\mathbf{a}) \|$ still being attained at the single point $x = 0$, and so the conditions of Theorem 3 are satisfied, showing that $\mathbf{a}^*$ gives a local best approximation. Notice that for this example

$$F^* = \{\mathbf{s} \in R^2, \|\mathbf{s}\| = 1, s_1 = 0\}$$

and so (2.2) of Lemma 4 is not satisfied.

EXAMPLE 2.

$$n = 3, \qquad l_1 = 1, \qquad l_2 = x, \qquad l_3 = x^2, \qquad m = \tfrac{1}{4}x^4 + \beta x^2.$$

Let $\mathbf{a}^* = (1, 0, -1)^T$. As in the previous example

$$V(l^*) = \{\delta(0)\}$$

and also

$$\begin{aligned} V(r^*) &= \operatorname{conv}\{\delta(-1), \delta(1)\} && \beta \leqslant -\tfrac{3}{2}, \\ &= \operatorname{conv}\{\delta(-\sqrt{-2-2\beta}), \delta(\sqrt{-2-2\beta})\} && -\tfrac{3}{2} \leqslant \beta \leqslant -1, \\ &= \{\delta(0)\} && -1 \leqslant \beta < \tfrac{3}{4}, \\ &= \operatorname{conv}\{-\delta(-1), \delta(0), -\delta(1)\} && \beta = \tfrac{3}{4}, \\ &= \operatorname{conv}\{-\delta(-1), -\delta(1)\} && \beta > \tfrac{3}{4}. \end{aligned}$$

Necessary conditions are satisfied for $-1 \leqslant \beta \leqslant \frac{3}{4}$ by taking $v = \delta(0)$, $\lambda = -1$. The set F^* is given by $\{\mathbf{s} \in R^3, \|\mathbf{s}\| = 1, s_1 = 0\}$ and so (2.2) of Lemma 4 is not satisfied for any of these values of β. In addition condition (ii) of Theorem 3 is not satisfied. However, when $-1 \leqslant \beta < \frac{3}{4}$, both $\|r(\mathbf{a})\|$ and $\|l(\mathbf{a})\|$ are differentiable functions at $\mathbf{a} = \mathbf{a}^*$, and thus the possibility exists of further information being provided by second-order conditions. Now if $-1 < \beta < \frac{3}{4}$, any perturbation of $\mathbf{a}^*$ will result in $\|r(\mathbf{a})\|$ and $\|l(\mathbf{a})\|$ both still being attained at the single points ξ and η respectively, where the derivatives with respect to x of the normed functions vanish. Thus derivatives in a neighbourhood of $\mathbf{a}^*$ are given by

$$\nabla \|r(\mathbf{a})\| = (1, \xi, \xi^2)^T,$$

where ξ must satisfy

$$a_2 + 2a_3\xi - \xi^3 - 2\beta\xi = 0$$

and $\nabla \|l(\mathbf{a})\| = (1, \eta, \eta^2)^T$, where $\eta = -a_2/2a_3$. It follows that the appropriate functions are twice differentiable in a neighbourhood of $\mathbf{a}^*$ and so Theorem 5 may be applied.

Now

$$\nabla^2\mathscr{L}(\mathbf{a}^*) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\beta/(2\beta+2) & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and so

$$\mathbf{s}^T\nabla^2\mathscr{L}(\boldsymbol{a}^*)\,\mathbf{s} = -\frac{\beta}{2\beta+2}\,s_2^{\,2}.$$

Thus the conditions of Theorem 5(i) are not satisfied if $\beta > 0$, and so for $0 < \beta < \frac{3}{4}$, l^* is not a local best approximation. Now let $\mathbf{a}^* = (-\frac{3}{4}, 0, \frac{7}{4})^T$. Then for $\beta \in [-1, \frac{3}{4}]$

$$\begin{aligned} V(l^*) &= \operatorname{conv}\{\delta(-1), \delta(1)\}; \\ V(r^*) &= \operatorname{conv}\{\delta(-1), \delta(1)\} && -1 \leqslant \beta < 0, \\ &= \operatorname{conv}\{\delta(-1), -\delta(0), \delta(1)\} && \beta = 0, \\ &= \{\delta(0)\} && 0 < \beta \leqslant \tfrac{3}{4}. \end{aligned}$$

Thus necessary conditions are satisfied for $-1 \leqslant \beta \leqslant 0$ by taking $v = w$, $\lambda = -1$ for any $w \in V(l^*)$. Also, since any perturbation of $\mathbf{a}^*$ results in $\| l(\mathbf{a})\|$ being attained at $x = -1$ and/or $x = +1$, Theorem 3 applies, showing that $\mathbf{a}^*$ is a local solution to (1.3) for these values of β.

References

1. B. L. Chalmers and G. D. Taylor, Uniform approximation with constraints, preprint, 1977.
2. A. V. Fiacco and G. P. McCormick, "Nonlinear Programming," Wiley, New York, 1968.
3. R. Fletcher and G. A. Watson, "First and Second Order Conditions for a Class of Nondifferentiable Optimization Problems," Dundee University Report NA/28 (1978).
4. R. B. Holmes, "Geometric Functional Analysis and its Applications," Springer-Verlag, New York, 1975.
5. R. T. Rockafellar, "Convex Analysis," Princeton Univ. Press, Princeton, N.J., 1970.
6. M. D. Ross and G. G. Belford, Characterisation of best Chebyshev approximations with prescribed norm, *J. Approximation Theory* **16** (1976), 315–328.
7. G. D. Taylor, Uniform approximation with side conditions, *in* "Approximation Theory" (G. G. Lorentz, Ed.), pp. 495–503, Academic Press, New York, 1973.
8. G. A. Watson, A class of programming problems whose objective function contains a norm, *J. Approximation Theory* **23** (1978), 401–411.
9. G. A. Watson, On a class of methods for nonlinear approximation problems, *in* "Proceedings, Durham Conference on Multivariate Approximation" (D. C. Hanscomb, Ed.), Academic Press, New York, 1978.