NEAREST POINTS IN CONVEX SETS IN NORMED LINEAR SPACES

Simon Fitzpatrick and John Giles

Abstract Given a norm-one linear functional u^* on a normed linear space E we give necessary and sufficient conditions for the separation by u^* of a convex set C from $B[x,d_C(x)]$ to imply the existence, or existence and uniqueness, of a nearest point in C to x.

1980 Mathematics Subject Classification Numbers (1985 Revision): 41A65, 46B20.

1. Introduction. Consider a normed linear space E and a nonempty closed subset K of E whose distance function d_K is defined by $d_K(x) := \inf\{\|x-z\| \mid z\in K\}$. We say that $x\in E\backslash K$ has a *nearest point* $p(x)$ in K provided $\|x - p(x)\| = d_K(x)$. If $x\in E\backslash K$ has a nearest point $x + d_C(x)u$ in a closed set K then u is a unit vector such that the one-sided directional derivative

$$d_K^+(x;u) := \lim_{t\searrow 0} t^{-1}[d_K(x+tu) - d_K(x)] = -1.$$

Recently, Fitzpatrick [Fi] has considered the converse problem; that is, if $x\in E\backslash K$ and $u\in S_E$ have $d_K^+(x;u) = -1$, under what circumstances must there be a nearest point in K to x? We consider here the particular case where K is convex.

Let C be a nonempty closed convex subset of E. It is clear that if $x\in E\backslash C$ has a nearest point $p(x)$ in C then there exists a support functional $u^*\in S_{E^*}$ to the closed ball $B[x,d_C(x)]$ at $p(x)$ which separates $B[x,d_C(x)]$ from C. In this paper we consider the converse, getting conditions on u^* under which if u^* separates $B[x,d_C(x)]$ from C then there must be a nearest point in C to x. We also explore conditions under which the existence and uniqueness of such nearest points is guaranteed.

The *duality mapping* on a normed linear space E is the multi-valued function defined by $J(x) := \{x^*\in E^* \mid \langle x^*,x\rangle = \|x\|^2 = \|x^*\|^2\}$. We let S_E and B_E denote the unit sphere and the closed unit ball of E and write $B[x,r] := x + rB_E$ for $x \in E$ and $r > 0$. We denote by $\hat{x}$ the canonical image of $x\in E$ in E^{**}.

2. Nearest points in convex sets. A nonempty closed set is convex if and only if its distance function is convex. We begin by giving a charaterization of the subdifferential of d_C for a convex set C and getting some extra information when $d_C^+(x;u) = -1$.

Proposition 1. Let C be a closed convex subset of a normed linear space E and $x \in E \setminus C$. Then $u^* \in \partial d_C(x)$ if and only if $u^* \in S_{E^*}$ and u^* separates C from $B[x, d_C(x)]$ with $\langle u^*, y\rangle \geq \langle u^*, c\rangle$ for all $c \in C$ and $y \in B[x, d_C(x)]$. If $u^* \in \partial d_C(x)$ and $d_C^+(x;u) = -1$ for some $u \in S_E$ then $-u^* \in Ju$.

Proof If $u^* \in \partial d_C(x)$ and $c \in C$ then $\langle u^*, c-x\rangle \leq d_C(c) - d_C(x) = -d_C(x)$. Taking $c \in C$ with $\|c-x\|$ close to $d_C(x)$ we have $\|u^*\| \geq 1$. But $\|u^*\| \leq 1$ as d_C has Lipschitz constant 1. Now if $c \in C$ then $\langle u^*, c\rangle \leq d_C(c) - d_C(x) + \langle u^*, x\rangle = -d_C(x) + \langle u^*, x\rangle$ which is the infimum of u^* on $B[x, d_C(x)]$ and u^* separates as required.

Conversely, if $u^* \in S_{E^*}$ and u^* separates C from $B[x, d_C(x)]$ with $\langle u^*, y\rangle \geq \langle u^*, c\rangle$ for all $c \in C$ and $y \in B[x, d_C(x)]$ then for $y \in B[x, d_C(x)]$ we have $d_C(y) \geq \sup\{\langle u^*, y-c\rangle \mid c \in C\} = \langle u^*, y-x\rangle + \sup\{\langle u^*, x-c\rangle \mid c \in C\} \geq \langle u^*, y-x\rangle + d_C(x)$ so that $u^* \in \partial d_C(x)$.

If we also have $d_C^+(x;u) = -1$ for some $u \in S_E$ then $\langle u^*, u\rangle \leq d_C^+(x;u) = -1$ and $\|u\| = 1 = \|u^*\|$ so $\langle u^*, u\rangle = -1$ and $-u^* \in Ju$. ♦

Theorem 2. Let E be a normed linear space and C a closed convex subset of E. If $x \in E \setminus C$ and $u^* \in S_{E^*}$ separates C from $B[x, d_C(x)]$ and if every sequence x_n from S_E with $\langle u^*, x_n\rangle \to 1$ has a weak cluster point then there is a nearest point in C to x.

Proof Let y_n be a minimizing sequence in C for x. We have, replacing u^* by $-u^*$ if necessary, $u^* \in \partial d_C(x)$ by Proposition 1. Thus $\langle u^*, y_n - x\rangle \leq d_C(y_n) - d_C(x) = -d_C(x)$ so that $\langle u^*, -y_n + x\rangle / \|y_n - x\| \to 1$. Thus $(x-y_n)/\|x-y_n\|$ has a weak cluster point and therefore y_n has a weak cluster point y. Since C is weakly closed we have $y \in C$ and we have $d_C(x) \leq \|x-y\| \leq \lim_n \|x-y_n\| = d_C(x)$ by weak lower semicontinuity of the norm, so y is a nearest point. ♦

The next lemma is basic in developing the converse of Theorem 2.

Lemma 3. Suppose $x_n \in S_E$ have no weakly convergent subsequence and $\lambda_n \to 1$. Then there is a subsequence y_j of $\lambda_n x_n$ such that

$$\bigcap_{m=1}^{\infty} \overline{\text{conv}}\left\{y_j \mid j>m\right\} = \emptyset.$$

Proof Without loss of generality E is separable. Then we may take a subsequence y_j of $\lambda_n x_n$ such that $\hat{y}_j$ is weak* convergent to some $F \in E^{**}$. Now we have

$$\bigcap_{m=1}^{\infty} \text{weak*}\overline{\text{conv}}\left\{\hat{y}_j \mid j>m\right\} = \{F\}$$

and since $F \in E^{**} \backslash \hat{E}$ our result follows. ♦

Theorem 4. Suppose $x_n \in S_E$ have no weakly convergent subsequence and let $u^* \in S_{E^*}$ with $\langle u^*, x_n\rangle \to 1$. Then there is a closed convex subset C of E and a point $x \in E \backslash C$ such that x has no nearest point in C and u^* separates C from $B[x, d_C(x)]$. If $u \in S_E$ and $-u^* \in Ju$ then we can also arrange to have $d_C^+(x;u) = -1$.

Proof Let $\lambda_n := (1+1/n)/\langle u^*, x_n\rangle$ and apply Lemma 3 to get a sequence y_n such that $\langle u^*, y_n\rangle \searrow 1$, $\|y_n\| \searrow 1$ and

$$\bigcap_{m=1}^{\infty} \overline{\text{conv}}\left\{y_j \mid j>m\right\} = \emptyset.$$

Take x such that $\langle u^*, x\rangle = \|x\|$ (it always is possible to take $x=0$) and $C := \overline{\text{conv}}\,\{-y_n \mid n \in \mathbb{N}\}$. We have $d_C(x) \le \|x\| + 1$ since $\|y_n\| \to 1$. On the other hand for $c \in C$ we have

$$\|x-c\| \ge \langle u^*, x - c\rangle \ge \inf\{\langle u^*, x + y_n\rangle \mid n \in \mathbb{N}\} = \|x\| + 1$$

which shows that $d_C(x) = \|x\| + 1$ and that

$$\sup\{\langle u^*, c\rangle \mid c \in C\} = \inf\{\langle u^*, y\rangle \mid y \in B[x, d_C(x)]\}.$$

Suppose z is a nearest point in C to x. Thus $\|x - z\| = \|x\| + 1$ and so $\langle u^*, x - z\rangle \le \|x\| + 1$. However $z \in \overline{\text{conv}}\,\{-y_n \mid n \in \mathbb{N}\}$ and $\langle u^*, y_n\rangle = 1 + 1/n$ so that $\langle u^*, x + y_n\rangle = \|x\| + 1 + 1/n$. Thus for each $m \in \mathbb{N}$ we have $z \in \overline{\text{conv}}\,\{-y_n \mid n>m\}$ and therefore

$$\bigcap_{m=1}^{\infty} \overline{\text{conv}}\left\{y_j \mid j>m\right\} \ne \emptyset$$

which is a contradiction.

Now let $-u^*\in Ju$ so we may take $x_s := (s-1)u$ in the calculations above where $0 \le s < 1$. Thus $d_C(x_s) = 1 + \|x_s\| = 1 + (1-s) = 2-s$ and we have $x_s+tu = x_{s+t}$ so $\lim_{t\searrow 0} t^{-1}[d_C(x_s+tu) - d_C(x_s)] = -1$. ♦

The following example illustrates Theorem 4 and shows that the converse to the last statement in Proposition 1 is not generally true.

Example 5. In $E := c_0$ let e_i denote the usual basis vectors and take $x := -e_1$, $u := e_1$ and $C := \overline{\text{conv}}\,\{(1+1/n)e_1 + e_2 + \ldots + e_n \mid n\in\mathbb{N}, n>1\}$. Then d_C has directional derivative -1 at $-e_1$ in the direction e_1 but $-e_1$ has no nearest point in C. At 0 d_C has one-sided directional derivative 0 in the direction e_1 but e_1^* separates C from $B[0,d_C(0)]$ and 0 has no nearest point in C. ♦

Combining these results we obtain the following characterizations.

Theorem 6. Let $u^*\in S_{E^*}$ on a normed linear space E. Every sequence x_n from S_E with $\langle u^*,x_n\rangle \to 1$ has a weak cluster point if and only if for every closed convex subset C of E and $x\in E\backslash C$ such that u^* separates C from $B[x,d_C(x)]$ there is a nearest point in C to x. ♦

For norm-attaining functionals we have an extra equivalent condition.

Theorem 7. Let E be a normed linear space and $u\in S_E$ and $u^*\in Ju$. The following are equivalent.

(i) Every sequence x_n from S_E with $\langle u^*,x_n\rangle \to 1$ has a weak cluster point.

(ii) For every closed convex subset C of E and $x\in E\backslash C$ such that u^* separates C from $B[x,d_C(x)]$ there is a nearest point in C to x.

(iii) For every closed convex subset C of E and $x\in E\backslash C$ and $v\in S_E$ such that $u^*\in\partial d_C(x)$ and $d_C^+(x;v) = -1$ there is a nearest point in C to x. ♦

3. Existence and uniqueness Now we turn to the question of uniqueness of nearest points in convex sets. If u^* does not attain its norm and u^* separates C from $B[x,d_C(x)]$ then there is no nearest point in C to x and the question of uniqueness does not arise.

Proposition 8. Let E be a normed linear space and let $u^* \in S_{E^*}$ attain its norm. Then u^* exposes B_E if and only if for every closed convex subset C of E and $x \in E \setminus C$ such that u^* separates C from $B[x,d_C(x)]$ there is at most one nearest point in C to x.

Proof If y and z are nearest points in C to x then $\|x-y\| = \|x-z\| = d_C(x)$ and $\langle u^*,x-y\rangle = \langle u^*,x-z\rangle = \pm d_C(x)$. Since u^* exposes $B[x,d_C(x)]$ we have $x-y = x-z$ and $y=z$ so the nearest point is unique if it exists.

Conversely suppose that u^* does not expose B_E. Thus there are y and z in S_E such that $y \neq z$ and $\langle u^*,y\rangle = \langle u^*,z\rangle = 1$. Let $C := \{ty + (1-t)z \mid 0 \leq t \leq 1\}$. Then $\langle u^*,x\rangle = 1 = d_C(0)$ for all $x \in C$ so u^* separates C from $B[0,d_C(0)]$ and every point of C is a nearest point in C to 0. ♦

Now we can characterize those u^* for which separation by u^* forces unique nearest points to exist.

Theorem 9. Let E be a normed linear space and $u \in S_E$ and $u^* \in Ju$. The following are equivalent.

(i) Every sequence x_n from S_E with $\langle u^*,x_n\rangle \to 1$ converges weakly to u.

(ii) The functional u^* exposes $B_{E^{**}}$ at $\hat{u}$.

(iii) The norm on E^* has Gateaux derivative $\hat{u}$ at u^*.

(iv) For every closed convex subset C of E and $x \in E \setminus C$ such that $\langle u^*,y\rangle \geq \langle u^*,c\rangle$ for all $c \in C$ and $y \in B[x,d_C(x)]$ the point $x - d_C(x)u$ is the unique nearest point in C to x.

(v) For every closed convex subset C of E and $x \in E \setminus C$ such that u^* separates C from $B[x,d_C(x)]$ there is a unique nearest point in C to x.

(vi) For every closed convex subset C of E and $x \in E \setminus C$ such that $u^* \in \partial d_C(x)$ and $v \in S_E$ with $d_C^+(x;v) = -1$ the point $x+d_C(x)v$ is the unique nearest point in C to x.

Proof If (i) holds then Theorem 7 shows the existence of a nearest point z in C to x in (iv), (v) and (vi), and Proposition 8 shows that z is the unique nearest point as u^* exposes B_E. But then $\langle u^*,x-z\rangle = d_C(x) = \|x-z\|$ and in (iv) we see that $\langle u^*,x - (x - d_C(x)u)\rangle = \| x - (x - d_C(x)u) \| = d_C(x)$ so that $z = x - d_C(x)u$. Now (vi) follows by Proposition 1 since $u^* \in J(-v)$. So (i) implies (iv), (v) and (vi). However (iv) implies (v) trivially and (vi) implies (v) by Proposition 1.

If (v) holds then Proposition 8 shows that u^* exposes B_E and Corollary 7 shows that every sequence x_n from S_E with $\langle u^*,x_n\rangle \to 1$ has a weak cluster point, which must be u because u^* exposes B_E at u. Suppose $F \in E^{**}$ with $\|F\| = 1$ and $\langle F,u^*\rangle = \langle u^*,u\rangle = 1$. Let $x^* \in E^*$. By weak* density of B_E in $B_{E^{**}}$ there are $x_n \in B_E$ such that $\langle x^*,x_n\rangle \to \langle F,x^*\rangle$ and $\langle u^*,x_n\rangle \to \langle F,u^*\rangle = 1$. Since u must be a weak cluster point of x_n we have $\langle x^*,u\rangle = \langle F,x^*\rangle$. Thus $F = \hat{u}$ as x^* was arbitrary, and u^* exposes $B_{E^{**}}$ at $\hat{u}$, giving (ii).

It is well known (see [Gi], p.197) that (iii) is equivalent to (ii). Finally suppose (ii) holds and let $x_n \in S_E$ with $\langle u^*,x_n\rangle \to 1$. We see that every weak* cluster point F of $\hat{x}_n$ has $\langle F,u^*\rangle = 1$ and $F \in B_{E^{**}}$ so that since u^* exposes $B_{E^{**}}$ at $\hat{u}$ we have $F = \hat{u}$ and $\hat{x}_n$ converges weak* to $\hat{u}$. However that means x_n converges weakly to u, yielding (i). ♦

4. Remarks Our theme in exploring the problem of the existence and uniqueness of nearest points in convex sets has concerned the nature of separating linear functionals. Much of the previous literature on nearest points in convex sets followed from the papers of Garkavi, but his work concerned quite different themes to ours. In [Ga2] Garkavi provided a necessary and sufficient condition for a point of a convex set to be a nearest point to a given point outside the set. Earlier, in [Ga1] he investigated the idea of embedding a convex set in the second dual and finding nearest points in the weak* closure of the set. We come closest to this idea when exploring the condition given in Theorem 9(ii).

References.

[Fi] Simon Fitzpatrick, "Nearest points to closed sets and directional derivatives of distance functions," Bull. Australian Math. Soc., to appear.

[Ga1] A. L. Garkavi, "Duality theory for approximation by elements of convex sets," Uspehi Mat. Nauk **16**(1961), 141-145.

[Ga2] A. L. Garkavi, "A criterion for the element of best approximation," Sibirsk. Mat. Ž. **5**(1964), 472-476.

[Gi] J. R. Giles, "Convex Analysis with Application in Differentiation of Convex Functions," Research Notes in Mathematics No. 58, Pitman, London, 1982.

Department of Mathematics and Statistics
University of Auckland,
Private bag, Auckland,
New Zealand.

Department of Mathematics
University of Newcastle
New South Wales 2308
Australia.