THE UNIQUE MINIMAL DUAL REPRESENTATION OF A CONVEX FUNCTION

HALUK ERGIN AND TODD SARVER

Abstract. Suppose (i) X is a separable Banach space, (ii) C is a convex subset of X that is a Baire space (when endowed with the relative topology) such that $\mathrm{aff}(C)$ is dense in X, and (iii) $f : C \to \mathbb{R}$ is locally Lipschitz continuous and convex. The Fenchel-Moreau duality can be stated as

$$f(x) = \max_{x^* \in \mathcal{M}} \big[\langle x, x^*\rangle - f^*(x^*)\big],$$

for all $x \in C$, where f^* denotes the Fenchel conjugate of f and $\mathcal{M} = X^*$. We show that, under assumptions (i)–(iii), there is a unique minimal weak*-closed subset $\mathcal{M}_f$ of X^* for which the above duality holds.

1. Introduction

Throughout, let X denote a real Banach space. Suppose $C \subset X$ and $f : C \to \mathbb{R}$. The *conjugate* (or *Fenchel conjugate*) of f is the function $f^* : X^* \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^*(x^*) = \sup_{x \in C} \big[\langle x, x^*\rangle - f(x)\big].$$

When f is a convex function, there is an important duality between f and f^* known as the Fenchel-Moreau theorem.[1] We next present a slight variation of this classic result when f is locally Lipschitz continuous. We relegate the proofs and the definitions of certain standard concepts to Section 2.

Lemma 1.1. *Suppose $C \subset X$ is convex and $f : C \to \mathbb{R}$ is locally Lipschitz continuous and convex. Then,*

$$f(x) = \max_{x^* \in X^*} \big[\langle x, x^*\rangle - f^*(x^*)\big] \tag{1.1}$$

for all $x \in C$.

Our main result shows that, under suitable assumptions on X, C, and f, there is a unique minimal weak*-closed subset of X^* for which Equation (1.1) holds. As we will explain in detail momentarily at the end of the Introduction, our result is motivated by recent work in theoretical economics. To construct this minimal set, we first introduce some necessary definitions.

Let C_f denote the set of all $x \in C$ for which the subdifferential of f at x is a singleton:

$$C_f = \{x \in C : \partial f(x) \text{ is a singleton}\}.$$

Date: March 9, 2010.

2000 *Mathematics Subject Classification.* Primary 46N10, 49N15; Secondary 46T20, 52A41.

Key words and phrases. Fenchel-Moreau duality, Mazur's theorem, singleton subdifferential.

We thank Kim Border, Massimo Marinacci, and a referee for helpful comments and suggestions.

[1] The standard version of this theorem states that if $f : X \to \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous and convex, then $f(x) = f^{**}(x) \equiv \sup_{x^* \in X^*}[\langle x, x^*\rangle - f^*(x^*)]$. See, e.g., Proposition 1 in [3, p97].

Let $\mathcal{N}_f$ denote the set of functionals contained in the subdifferential of f at some $x \in C_f$:

$$\mathcal{N}_f = \{x^* \in X^* : x^* \in \partial f(x) \text{ for some } x \in C_f\}.$$

Finally, let $\mathcal{M}_f$ denote the closure of $\mathcal{N}_f$ in the weak* topology:

$$\mathcal{M}_f = \overline{\mathcal{N}_f}.$$

We are now ready to state our main result.

Theorem 1.2. *Suppose (i) X is a separable Banach space, (ii) C is a convex subset of X that is a Baire space (when endowed with the relative topology) such that $\operatorname{aff}(C)$ is dense in X, and (iii) $f : C \to \mathbb{R}$ is locally Lipschitz continuous and convex. Then, for any weak*-closed $\mathcal{M} \subset X^*$, the following are equivalent:*

(1) $\mathcal{M}_f \subset \mathcal{M}$.
(2) *For all $x \in C$, the maximization problem*

$$\max_{x^* \in \mathcal{M}} \left[\langle x, x^* \rangle - f^*(x^*)\right],$$

has a solution and the maximum value is equal to $f(x)$.

We present the proof of Theorem 1.2 in Section 2. In Section 3, we provide counterexamples to illustrate why some of the assumptions in Theorem 1.2 cannot be relaxed. Example 3.1 shows that if the local Lipschitz continuity assumption is weakened to continuity in Theorem 1.2, then it is possible to have $\mathcal{M}_f = \emptyset$. Example 3.2 shows that Theorem 1.2 fails to hold if the maximum in part (2) is not assumed to exist and maximum is replaced with supremum. However, in the special case when f is Lipschitz continuous, $\mathcal{M}_f$ is compact and the existence of the maximum in (2) can be guaranteed by restricting attention to compact $\mathcal{M}$. The following corollary formalizes this observation:

Corollary 1.3. *Suppose X, C, and f satisfy (i)-(iii) in Theorem 1.2 and f is Lipschitz continuous. Then, $\mathcal{M}_f$ is weak* compact, and for any weak*-compact $\mathcal{M} \subset X^*$,*

$$\mathcal{M}_f \subset \mathcal{M} \iff f(x) = \max_{x^* \in \mathcal{M}} \left[\langle x, x^* \rangle - f^*(x^*)\right] \quad \forall x \in C.$$

Our motivation for these results comes from decision theory in theoretical economics, where elements of C are interpreted as choice objects. The primitive is a binary relation over C interpreted as an individual's preferences over C. In several applications of interest (see, e.g., [4, 5, 6]), such a binary relation can be represented by a function f satisfying the conditions of Theorem 1.2, i.e., x is preferred to y if and only if $f(x) > f(y)$. In these applications, the duality formula may be interpreted as the individual's anticipation that after she chooses x, an unobservable costly action x^* will be selected. Her payoff from action x^* is given by $\langle x, x^* \rangle$ and the cost of the action is given by $f^*(x^*)$. The uniqueness of the minimal subset of X^* established in Theorem 1.2 identifies the set of relevant available actions from the binary relation on C. Identifying the minimal relevant set of maximizers is also potentially useful in optimization theory where the value function f results from a maximization as in Equation (1.1).

2. Proof of Theorem 1.2

We start by briefly stating the definitions of certain standard concepts that will be used frequently in the sequel.

Definition 2.1. For $C \subset X$, a function $f : C \to \mathbb{R}$ is said to be *Lipschitz continuous* if there is some real number K such that $|f(x)-f(y)| \leq K\|x-y\|$ for every $x, y \in C$. The number K is called a *Lipschitz constant* of f. A function f is said to be *locally Lipschitz continuous* if for every $x \in C$, there exists $\varepsilon > 0$ such that f is Lipschitz continuous on $B_\varepsilon(x) \cap C = \{y \in C : \|y - x\| < \varepsilon\}$.

Definition 2.2. Suppose $C \subset X$ and $f : C \to \mathbb{R}$. For $x \in C$, the *subdifferential* of f at x is defined to be

$$\partial f(x) = \{x^* \in X^* : \langle y - x, x^* \rangle \leq f(y) - f(x) \text{ for all } y \in C\}.$$

Definition 2.3. The *affine hull* of a set $C \subset X$, denoted $\operatorname{aff}(C)$, is defined to be the smallest affine subspace of X that contains C.

The proof of Theorem 1.2 relies on a key intermediate result generalizing the theorem of Mazur (1933) on the generic Gâteaux differentiability of continuous convex functions. Mazur [10] showed that if X is a separable Banach space and $f : C \to \mathbb{R}$ is a continuous convex function defined on a convex open subset C of X, then the set of points x where f is Gâteaux differentiable is a dense G_δ set in C.[2] We next extend Mazur's theorem by replacing the assumption that C is open with the weaker assumptions that C is a Baire space (when endowed with the relative topology) and that the affine hull of C is dense in X.

Theorem 2.4. *Suppose X, C, and f satisfy (i)-(iii) in Theorem 1.2. Then, the set of points x where $\partial f(x)$ is a singleton is a dense G_δ set (in the relative topology) in C.*

Note that Mazur's theorem is a special case of Theorem 2.4. First, if C is an open subset of X, then C is a Baire space and $\operatorname{aff}(C) = X$. Second, any continuous convex function f defined on an open set C is locally Lipschitz continuous (see [11, Proposition 1.6]). Therefore, if C is open, then our continuity assumption coincides with that of Mazur's theorem. Finally, it is a standard result that a continuous convex function f defined on an open set C is Gâteaux differentiable at a point x if and only if the subdifferential $\partial f(x)$ is a singleton set (see [11, Proposition 1.8]). Thus, if C is open, then the conclusion of Theorem 2.4 also coincides with the conclusion of Mazur's theorem.

The equivalence of assumptions described in the previous paragraph need not hold if C is not open, and we provide two examples in Section 3 to illustrate. We show in Example 3.1 that there exists a set C satisfying the assumptions of Theorem 2.4 and a function $f : C \to \mathbb{R}$ that is both continuous and linear on C such that $\partial f(x)$ is not a singleton for any $x \in C$. By Theorem 2.4, this implies that f is not locally Lipschitz continuous. In Example 3.3, we show that there exists a set C and a function $f : C \to \mathbb{R}$ satisfying the assumptions of Theorem 2.4 such that f is not Gâteaux differentiable at any $x \in C$.

We will present a direct proof of Theorem 2.4 which follows a similar approach to the proof of Mazur's theorem found in [11]. Theorem 2.4 can be shown to follow

[2] For a textbook treatment, see [11, Theorem 1.20]. An equivalent characterization in terms of closed convex sets and smooth points can be found in [9, p171].

indirectly from results in [1, 2, 7, 8, 12, 13]. We present the direct proof because it is self-contained and constructive.

We begin by establishing that the subdifferential of a Lipschitz continuous and convex function is nonempty at every point.

Lemma 2.5. *Suppose C is a convex subset of a Banach space X. If $f : C \to \mathbb{R}$ is Lipschitz continuous and convex, then $\partial f(x) \neq \emptyset$ for all $x \in C$. In particular, if $K \geq 0$ is a Lipschitz constant of f, then for all $x \in C$ there exists $x^* \in \partial f(x)$ with $\|x^*\| \leq K$.*

Proof. We only outline the proof of this lemma since it is standard. Fix any $x \in C$, and define

$$H(x) = \{(y, t) \in X \times \mathbb{R} : t < f(x) - K\|y - x\|\}.$$

Then, $H(x)$ and $\text{epi}(f)$ are disjoint, where $\text{epi}(f)$ denotes the epigraph of f:

$$\text{epi}(f) = \{(y, t) \in C \times \mathbb{R} : t \geq f(y)\}.$$

Since $H(x)$ has a nonempty interior, it can be shown there exists $x^* \in X^*$ such that $(x^*, -1) \in X^* \times \mathbb{R}$ separates $H(x)$ and $\text{epi}(f)$. This implies that $x^* \in \partial f(x)$ and $\|x^*\| \leq K$. $\square$

By definition, for every point in the domain of a locally Lipschitz continuous function, there exists a neighborhood on which the function is Lipschitz continuous. Therefore, the following lemma allows the preceding result to be applied to locally Lipschitz functions. We omit the straightforward proof.

Lemma 2.6. *Suppose C is a convex subset of a Banach space X, and fix any $x \in C$ and $\varepsilon > 0$. Then, $\partial f(y) = \partial f|_{B_\varepsilon(x)\cap C}(y)$ for all $y \in B_\varepsilon(x) \cap C$.*

Suppose X, C, and f satisfy (i)-(iii) in Theorem 1.2. Note that for any $x, y \in C$, we have $\text{span}(C - x) = \text{span}(C - y)$. In addition, since $\text{aff}(C)$ is dense in X, it follows that $\text{span}(C - y) = \text{aff}(C) - y$ is also dense in X. Since any subset of a separable Banach space is separable, $\text{span}(C - y)$ is separable for any $y \in C$. Let $\{x_n\} \subset \text{span}(C - y)$ be a sequence which is dense in $\text{span}(C - y)$ and hence also dense in X. For each $K, m, n \in \mathbb{N}$, let $A_{K,m,n}$ denote the set of all $x \in C$ for which there exist $x^*, y^* \in \partial f(x)$ such that

$$\|x^*\|, \|y^*\| \leq K \text{ and } \langle x_n, x^* - y^* \rangle \geq \tfrac{1}{m}.$$

The following lemmas establish the key properties of $A_{K,m,n}$ that will be needed for our proof of Theorem 2.4.

Lemma 2.7. *Suppose X, C, and f satisfy (i)-(iii) in Theorem 1.2. Then, the set of $x \in C$ for which $\partial f(x)$ is a singleton is $\bigcap_{K,m,n}(C \setminus A_{K,m,n})$.*

Proof. Clearly, if $\partial f(x)$ is a singleton, then $x \in \bigcap_{K,m,n}(C \setminus A_{K,m,n})$. To prove the converse, we will show that if $\partial f(x)$ is not a singleton for $x \in C$, then $x \in A_{K,m,n}$ for some $K, m, n \in \mathbb{N}$. We first show that $\partial f(x) \neq \emptyset$ for all $x \in C$. To see this, fix any $x \in C$. Since f is locally Lipschitz continuous, there exists $\varepsilon > 0$ such that f is Lipschitz continuous on $B_\varepsilon(x) \cap C$. Therefore, by Lemma 2.5, $\partial f|_{B_\varepsilon(x)\cap C}(x) \neq \emptyset$. By Lemma 2.6, this implies that $\partial f(x) \neq \emptyset$.

Suppose $\partial f(x)$ is not a singleton. Since $\partial f(x)$ is nonempty, there exist $x^*, y^* \in \partial f(x)$ such that $x^* \neq y^*$. Hence, there exists $y \in X$ such that $\langle y, x^* - y^* \rangle > 0$. Since $\{x_n\}$ is dense in X, by the continuity of $x^* - y^*$, there exists $n \in \mathbb{N}$ such that

$\langle x_n, x^* - y^* \rangle > 0$. Thus, there exists $m \in \mathbb{N}$ such that $\langle x_n, x^* - y^* \rangle \geq \frac{1}{m}$. Therefore, taking $K \in \mathbb{N}$ such that $\|x^*\|, \|y^*\| \leq K$, we have $x \in A_{K,m,n}$. □

Lemma 2.8. *Suppose X, C, and f satisfy (i)-(iii) in Theorem 1.2. Then, $A_{K,m,n}$ is a closed subset of C (in the relative topology) for any $K, m, n \in \mathbb{N}$.*

Proof. Consider any sequence $\{z_k\} \subset A_{K,m,n}$ such that $z_k \to z$ for some $z \in C$. We will show that $z \in A_{K,m,n}$. For each k, choose $x_k^*, y_k^* \in \partial f(z_k)$ such that $\|x_k^*\|, \|y_k^*\| \leq K$ and $\langle x_n, x_k^* - y_k^* \rangle \geq \frac{1}{m}$. Since $\{x^* \in X^* : \|x^*\| \leq K\}$ is weak* compact by Alaoglu's theorem and weak* metrizable by the separability of X, any sequence in this ball has a weak*-convergent subsequence. Thus, without loss of generality, we can assume there exist $x^*, y^* \in X^*$ with $\|x^*\|, \|y^*\| \leq K$ such that $x_k^* \xrightarrow{w^*} x^*$ and $y_k^* \xrightarrow{w^*} y^*$. Hence, by the norm-boundedness of the sequence $\{x_k^*\}$, the definition of the subdifferential, and the continuity of f, for any $y \in C$,

$$\langle y - z, x^* \rangle = \lim_k \langle y - z_k, x_k^* \rangle \leq \lim_k \big[f(y) - f(z_k) \big] = f(y) - f(z),$$

which implies $x^* \in \partial f(z)$.[3] A similar argument shows $y^* \in \partial f(z)$. Finally, since

$$\langle x_n, x^* - y^* \rangle = \lim_k \langle x_n, x_k^* - y_k^* \rangle \geq \tfrac{1}{m},$$

we have $z \in A_{K,m,n}$, and hence $A_{K,m,n}$ is relatively closed. □

Lemma 2.9. *Suppose X, C, and f satisfy (i)-(iii) in Theorem 1.2. Then, $C \setminus A_{K,m,n}$ is dense in C for any $K, m, n \in \mathbb{N}$.*

Proof. Since C is convex, it is straightforward to show that

$$\operatorname{aff}(C) = \{\lambda x + (1-\lambda) y : x, y \in C \text{ and } \lambda \in \mathbb{R}\}. \tag{2.1}$$

Consider arbitrary $K, m, n \in \mathbb{N}$ and $z \in C$. We will find a sequence $\{z_k\} \subset C \setminus A_{K,m,n}$ such that $z_k \to z$. Recall that $x_n \in \operatorname{span}(C - y)$ for any choice of $y \in C$. Thus, $z + x_n \in z + \operatorname{span}(C - z) = \operatorname{aff}(C)$, so by Equation (2.1), there exist $x, y \in C$ and $\lambda \in \mathbb{R}$ such that $\lambda x + (1-\lambda) y = z + x_n$. Let us first suppose $\lambda > 1$; we will consider the other cases shortly. Note that $\lambda > 1$ implies $0 < \frac{\lambda - 1}{\lambda} < 1$. Consider any sequence $\{a_k\} \subset (0, \frac{\lambda-1}{\lambda})$ such that $a_k \to 0$. Define a sequence $\{y_k\} \subset C$ by $y_k = a_k y + (1 - a_k) z$, and note that $y_k \to z$. We claim that for each $k \in \mathbb{N}$, $y_k + \frac{a_k}{\lambda - 1} x_n \in C$. To see this, note the following:

$$\begin{aligned} y_k + \tfrac{a_k}{\lambda-1} x_n &= a_k y + (1 - a_k) z + \tfrac{a_k}{\lambda - 1} \big(\lambda x + (1-\lambda) y - z \big) \\ &= \big(1 - \tfrac{a_k \lambda}{\lambda - 1} \big) z + \tfrac{a_k \lambda}{\lambda - 1} x. \end{aligned}$$

Since $0 < a_k < \frac{\lambda-1}{\lambda}$, we have $0 < \frac{a_k \lambda}{\lambda - 1} < 1$. Thus, $y_k + \frac{a_k}{\lambda-1} x_n$ is a convex combination of z and x, so it is an element of C. This is illustrated in Figure 1.

Consider any $k \in \mathbb{N}$. Because C is convex, we have $y_k + t x_n \in C$ for all $t \in (0, \frac{a_k}{\lambda - 1})$. Define a function $g : (0, \frac{a_k}{\lambda-1}) \to \mathbb{R}$ by $g(t) = f(y_k + t x_n)$, and note that g is convex. It is a standard result that a convex function on an open interval

[3] The first equality follows from a standard result: Let $K \geq 0$ and let $\{z_k\} \subset X$ and $\{z_k^*\} \subset X^*$ be sequences such that (i) $\|z_k^*\| \leq K$ for all k, and (ii) $z_k \to z$ and $z_k^* \xrightarrow{w^*} z^*$ for some $z \in X$ and $z^* \in X^*$. Then,

$$\begin{aligned} |\langle z_k, z_k^* \rangle - \langle z, z^* \rangle| &\leq |\langle z_k - z, z_k^* \rangle| + |\langle z, z_k^* - z^* \rangle| \leq \|z_k - z\| \|z_k^*\| + |\langle z, z_k^* - z^* \rangle| \\ &\leq \|z_k - z\| K + |\langle z, z_k^* - z^* \rangle| \to 0, \end{aligned}$$

so $\langle z_k, z_k^* \rangle \to \langle z, z^* \rangle$.

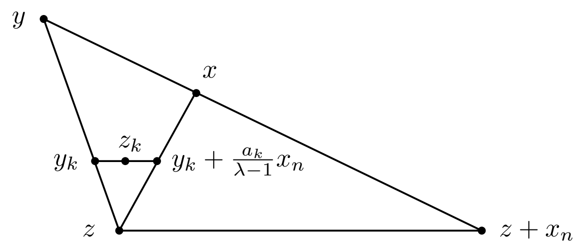

FIGURE 1. Construction of the sequence $\{z_k\}$

in $\mathbb{R}$ is differentiable for all but (at most) countably many points of this interval (see [11, Theorem 1.16]). Let $t_k \in (0, \frac{a_k}{\lambda-1})$ be such that $g'(t_k)$ exists, and let $z_k = y_k + t_k x_n$. If $x^* \in \partial f(z_k)$, then it is straightforward to verify that the linear mapping $t \mapsto t\langle x_n, x^*\rangle$ is in the subdifferential of g at t_k. Since g is differentiable at t_k, it can only have one element in its subdifferential at that point. Therefore, for any $x^*, y^* \in \partial f(z_k)$, we have $\langle x_n, x^*\rangle = \langle x_n, y^*\rangle$; hence, $z_k \in C \setminus A_{K,m,n}$. Finally, note that since $0 < t_k < \frac{a_k}{\lambda-1}$ and $a_k \to 0$, we have $t_k \to 0$. Therefore, $z_k = y_k + t_k x_n \to z$.

Above, we did restrict attention to the case of $\lambda > 1$. However, if $\lambda < 0$, then let $\lambda' = 1 - \lambda > 1$, $x' = y$, $y' = x$, and the analysis is the same as above. If $\lambda \in [0, 1]$, then note that $z + x_n \in C$. Similar to the preceding paragraph, for any $k \in \mathbb{N}$, define a function $g : (0, \frac{1}{k}) \to \mathbb{R}$ by $g(t) = f(z + tx_n)$. Let $t_k \in (0, \frac{1}{k})$ be such that $g'(t_k)$ exists, and let $z_k = z + t_k x_n$. Then, as argued above, we have $z_k \in C \setminus A_{K,m,n}$ for all $k \in \mathbb{N}$ and $z_k \to z$. $\square$

Proof of Theorem 2.4. By Lemma 2.7, the set of $x \in C$ for which $\partial f(x)$ is a singleton is $\bigcap_{K,m,n}(C \setminus A_{K,m,n})$. By Lemmas 2.8 and 2.9, for each $K, m, n \in \mathbb{N}$, $C \setminus A_{K,m,n}$ is open (in the relative topology) and dense in C. Since C is a Baire space, every countable intersection of open dense subsets of C is also dense. This completes the proof. $\square$

Lemma 2.10 summarizes certain properties of f^* that are useful in establishing the variation of the Fenchel-Moreau duality stated in Lemma 1.1. The proof of Lemma 2.10 is standard (see, e.g., [3]); it is therefore omitted.

Lemma 2.10. *Suppose $C \subset X$ and $f : C \to \mathbb{R}$. Then,*

(1) *f^* is lower semi-continuous in the weak* topology.*
(2) *$f(x) \geq \langle x, x^*\rangle - f^*(x^*)$ for all $x \in C$ and $x^* \in X^*$.*
(3) *$f(x) = \langle x, x^*\rangle - f^*(x^*)$ if and only if $x^* \in \partial f(x)$.*

We now prove the results stated in the Introduction.

Proof of Lemma 1.1. For any $x \in C$, there exists $\varepsilon > 0$ such that f is Lipschitz continuous on $B_\varepsilon(x) \cap C$. By Lemma 2.5, $\partial f|_{B_\varepsilon(x)\cap C}(x) \neq \emptyset$. By Lemma 2.6, this implies that $\partial f(x) \neq \emptyset$. Therefore, parts (2) and (3) in Lemma 2.10 imply Equation (1.1) for all $x \in C$. $\square$

Proof of Theorem 1.2. ($1 \Rightarrow 2$): Let $x \in C$ be arbitrary. By Theorem 2.4, C_f is dense in C, so there exists a sequence $\{x_k\} \subset C_f$ such that $x_k \to x$.[4] Since f is locally Lipschitz continuous, there exists $\varepsilon > 0$ such that f is Lipschitz continuous on $B_\varepsilon(x) \cap C$. Let $K \geq 0$ be a Lipschitz constant of $f|_{B_\varepsilon(x)\cap C}$. Without loss of generality assume that $x_k \in B_\varepsilon(x)$ for all k.

For each k, by Lemma 2.5, there exists $x_k^* \in \partial f|_{B_\varepsilon(x)\cap C}(x_k)$ such that $\|x_k^*\| \leq K$. By Lemma 2.6, $x_k^* \in \partial f|_{B_\varepsilon(x)\cap C}(x_k) = \partial f(x_k)$. Therefore, $\{x_k^*\} \subset \mathcal{M}_f \cap \{x^* \in X^* : \|x^*\| \leq K\}$, where the intersection is weak* compact and weak* metrizable since $\mathcal{M}_f$ is weak* closed, $\{x^* \in X^* : \|x^*\| \leq K\}$ is weak* compact by Alaoglu's theorem, and X is separable. Thus, $\{x_k^*\}$ has a convergent subsequence. Without loss of generality, suppose the sequence itself converges, so that $x_k^* \xrightarrow{w^*} x^*$ for some $x^* \in \mathcal{M}_f$. By the norm-boundedness of the sequence $\{x_k^*\}$, the definition of the subdifferential, and the continuity of f, for any $y \in C$,

$$\langle y - x, x^* \rangle = \lim_k \langle y - x_k, x_k^* \rangle \leq \lim_k \big[f(y) - f(x_k)\big] = f(y) - f(x),$$

which implies $x^* \in \partial f(x)$. Since $x \in C$ was arbitrary, we conclude that for all $x \in C$, there exists $x^* \in \mathcal{M}_f \subset \mathcal{M}$ such that $x^* \in \partial f(x)$. Then, by (2) and (3) in Lemma 2.10, we conclude that for all $x \in C$,

$$f(x) = \max_{x^* \in \mathcal{M}} \big[\langle x, x^* \rangle - f^*(x^*)\big].$$

($2 \Rightarrow 1$): Fix any $x \in C_f$. Since the maximization in part 2 is assumed to have a solution, there exists $x^* \in \mathcal{M}$ such that $f(x) = \langle x, x^* \rangle - f^*(x^*)$, which implies $x^* \in \partial f(x)$ by (3) in Lemma 2.10. However, $x \in C_f$ implies $\partial f(x) = \{x^*\}$, and hence $\partial f(x) \subset \mathcal{M}$. Since $x \in C_f$ was arbitrary, we have $\mathcal{N}_f \subset \mathcal{M}$. Because $\mathcal{M}$ is weak* closed, we have $\mathcal{M}_f = \overline{\mathcal{N}_f} \subset \mathcal{M}$. □

3. Examples

Let $X = l_1$, so $X^* = l_\infty$. Let $C = \{x \in l_1 : -1/i^3 \leq x_i \leq 1/i^3\}$. It is standard to verify that l_1 and C satisfy the assumptions of Theorem 1.2, but the interior of C is empty. Define the ith unit vector $e^i \in l_1$ by $e^i_i = 1$ and $e^i_j = 0$ for all $j \neq i$.

We first give an example of a continuous linear function $f : C \to \mathbb{R}$ such that $\partial f(x)$ is not a singleton for any $x \in C$. By Theorem 2.4, this function f cannot be locally Lipschitz continuous. In particular, Example 3.1 shows that the local Lipschitz continuity assumption cannot be dropped from Theorems 1.2 and 2.4.

Example 3.1. Let $f : C \to \mathbb{R}$ be defined by $f(x) = \sum_{i=1}^\infty i x_i$. This is well-defined since for any $x \in C$, $\sum_i |i x_i|$ is a series with positive terms that is bounded above by the convergent series $\sum_i 1/i^2$. We first show that f is continuous on C. Note that for any $I \in \mathbb{N}$ and $x, y \in C$,

$$\begin{aligned} |f(x) - f(y)| &\leq \sum_{i \leq I} |i x_i - i y_i| + \sum_{i > I} |i x_i - i y_i| \\ &\leq \sum_{i \leq I} I |x_i - y_i| + \sum_{i > I} 2/i^2 < I \|x - y\|_1 + 2/I, \end{aligned}$$

[4] If C were also assumed to be open, then one could apply Mazur's theorem here instead of Theorem 2.4. However, in a number of applications, such as [4, 5, 6], the domain C has an empty interior, yet it satisfies the assumptions of our Theorem 2.4.

since $\sum_{i>I} 1/i^2 < 1/I$. For any $\varepsilon > 0$, there is $I \in \mathbb{N}$ such that $2/I < \varepsilon$. From above, for any $x, y \in C$ satisfying $\|x-y\|_1 < (1/I)(\varepsilon - 2/I)$, we have $|f(x)-f(y)| < \varepsilon$. Therefore, f is continuous on C.

We now show that $\partial f(x)$ is not a singleton for any $x \in C$. Suppose to the contrary that $\partial f(x) = \{x^*\}$ for some $x^* \in l_\infty$. First, consider the case where $x_i > -1/i^3$ for all $i \in \mathbb{N}$. For all $i \in \mathbb{N}$, there exists $\lambda > 0$ such that $x - \lambda e^i \in C$. Then, $-\lambda i = f(x - \lambda e^i) - f(x) \geq \langle -\lambda e^i, x^* \rangle = -\lambda x_i^*$, implying $x_i^* \geq i$. Since the latter holds for all $i \in \mathbb{N}$, this contradicts $x^* \in l_\infty$.

Next, consider the case where $x = -1/i^3$ for some i. Define $y^* = x^* - e^i \in l_\infty$ and take any $y \in C$. Note that $\langle y - x, e^i \rangle = y_i + 1/i^3 \geq 0$. Therefore, since $x^* \in \partial f(x)$, we have

$$f(y) - f(x) \geq \langle y - x, x^* \rangle \geq \langle y - x, x^* \rangle - \langle y - x, e^i \rangle = \langle y - x, y^* \rangle.$$

Since the above equation holds for all $y \in C$, we have $y^* \in \partial f(x)$. Since $y^* \neq x^*$, this contradicts $\partial f(x) = \{x^*\}$.

To directly see that f is not Lipschitz continuous, fix any $i \in \mathbb{N}$, and take $\lambda > 0$ such that $\lambda e^i \in C$. Then, we have $f(\lambda e^i) - f(0) = i\lambda = i\|\lambda e^i - 0\|_1$. Therefore, f is not Lipschitz continuous. It is easy to see that local Lipschitz continuity and Lipschitz continuity are equivalent for a linear function on a convex subset of a normed linear space. Therefore, f is also not locally Lipschitz continuous.

We next give an example illustrating that if in part (2) of Theorem 1.2 we drop the assumption that the maximization problem has a solution and replace the maximum operator with the supremum operator, then (2) does not imply (1).

Example 3.2. Define $f : C \to \mathbb{R}$ by $f(x) = 0$ for all $x \in C$. Clearly, X, C, and f satisfy (i)–(iii) in Theorem 1.2. By Theorem 2.4, the set of points $x \in C$ where $\partial f(x)$ is a singleton is a dense subset of C. In fact, it is easy to verify that $C_f = \{x \in l_1 : -1/i^3 < x_i < 1/i^3\}$ and $\mathcal{N}_f = \mathcal{M}_f = \{0\}$.

For each $i \in \mathbb{N}$, define $x^{*i} = i^2 e^i$ and let $\mathcal{M} = \{x^{*i} : i \in \mathbb{N}\}$. Then, for $1 < \alpha < 2$ and $x = (i^{-\alpha}) \in l_1 \setminus C$, $\langle x, x^{*i} \rangle = i^{2-\alpha} \to \infty$. This implies that $\mathcal{M}$ is weak* closed. Note also that

$$f^*(x^{*i}) = \sup_{x \in C} \left[\langle x, x^{*i} \rangle - f(x)\right] = \frac{1}{i}$$

since the supremum is attained at $x = \frac{1}{i^3} e^i$. Then, for all $x \in C$,

$$f(x) = 0 = \sup_{i \in \mathbb{N}} \left[i^2 x_i - \frac{1}{i}\right] = \sup_{x^* \in \mathcal{M}} \left[\langle x, x^* \rangle - f^*(x)\right]$$

since $(i^2 x_i - 1/i) \nearrow 0$. However, $\mathcal{M}_f = \{0\}$ is not a subset of $\mathcal{M}$.

If a function is convex and continuous on an *open* and convex domain, then Phelps [11] shows that Gâteaux differentiability is equivalent to having a singleton subdifferential. We next give an example of a function $f : C \to \mathbb{R}$ such that X, C, and f satisfy (i)–(iii) in Theorem 1.2 such that f is not Gâteaux differentiable at any $x \in C$.

Example 3.3. Again, let $f(x) = 0$ for all $x \in C$. As noted in Example 3.2, $C_f = \{x \in l_1 : -1/i^3 < x_i < 1/i^3\}$ and $\mathcal{N}_f = \mathcal{M}_f = \{0\}$. Now, let $y = (1/i^2) \in l_1$. Then, for any $x \in C$ and $\lambda > 0$, taking $i > 2/\lambda$ gives $\lambda y_i = \lambda/i^2 > 2/i^3$. This implies that $x_i + \lambda y_i > -1/i^3 + 2/i^3 = 1/i^3$. Thus, $x + \lambda y \notin C$ for any $\lambda > 0$. Therefore, f cannot be Gâteaux differentiable at x.

References

[1] Borwein, J., Fitzpatrick, S., and P. Kenderov (1991): "Minimal Convex Uscos and Monotone Operators on Small Sets," *Canadian Journal of Mathematics*, 43, 461–476.

[2] Borwein, J. and R. Goebel (2003): "Notions of Relative Interior in Banach Spaces," *Journal of Mathematical Sciences*, 115, 2542-2553.

[3] Ekeland, I., and T. Turnbull (1983): *Infinite-Dimensional Optimization and Convexity*. Chicago: The University of Chicago Press.

[4] Epstein, L. G., M. Marinacci, and K. Seo (2007): "Coarse Contingencies and Ambiguity," *Theoretical Economics*, 2, 355-394.

[5] Ergin, H. and T. Sarver (2009): "A Subjective Model of Temporal Preferences," Mimeo, Northwestern University and Washington University in Saint Louis.

[6] Ergin, H. and T. Sarver (2010): "A Unique Costly Contemplation Representation," *Econometrica*, forthcoming.

[7] Fabian M. J. (1997): *Gâteaux Differentiability of Convex Functions and Topology*. New York: John Wiley & Sons.

[8] Floyd, E. E. , and V. L. Klee (1954), "A Characterization of Reflexivity by the Lattice of Closed Subspaces," *Proceedings of the American Mathematical Society*, 5, 655–661.

[9] Holmes, R. (1975): *Geometric Functional Analysis and Its Applications*. New York: Springer-Verlag.

[10] Mazur, S. (1933): "Über konvexe Mengen in linearen normierten Räumen," *Studia Mathematica*, 4, 70–84.

[11] Phelps, R. R. (1993): *Convex Functions, Monotone Operators, and Differentiability*. Berlin, Germany: Springer-Verlag.

[12] Verona, M. E. (1988): "More on the Differentiability of Convex Functions," *Proceedings of the American Mathematical Society*, 103, 137–140.

[13] Verona A., and M. E. Verona (1990): "Locally Efficient Monotone Operators," *Proceedings of the American Mathematical Society*, 109, 195–204.

Washington University in Saint Louis, Department of Economics, Campus Box 1208, Saint Louis, Missouri 63130

E-mail address: hergin@artsci.wustl.edu

Northwestern University, Department of Economics, 2001 Sheridan Road, Evanston, Illinois 60208

E-mail address: tsarver@northwestern.edu