SUBGRADIENTS OF A CONVEX FUNCTION OBTAINED FROM A DIRECTIONAL DERIVATIVE

PETER D. TAYLOR

Suppose R is a lower semicontinuous convex function on a Banach space E. A new result is obtained relating the directional derivatives of h and its subgradients: if l is a tangent line at some point z in graph h then a hyperplane can be found in $E \times R$ which supports epigraph h at a point close to z and almost contains l. This theorem is applied to get a formula for the directional derivative of h at a point in terms of the derivatives in the same direction of subgradients at nearby points. This formula is used to obtain several known results including the maximal monotonicity of the subdifferential of h and the uniqueness of h with a given subdifferential.

The main lemma takes a point z in a closed convex set C, and a bounded set X, all in a Banach space E, and gives conditions under which there exists a hyperplane which supports C at a point close to z and separates C and X.

A *proper convex function* on a real Banach space E is a function h on E with values in $(-\infty, +\infty)$, not identically $+\infty$, such that

$$h(\lambda x + (1 - \lambda)y) \leqq \lambda h(x) + (1 - \lambda)h(y)$$

for $x, y \in E$ and $0 \leqq \lambda \leqq 1$. A subgradient of h at $x \in E$ is an $s \in E^*$ such that

$$h(y) \geqq h(x) - s(y - x) \qquad \text{for all} \quad y \in E .$$

This says that $h(x)$ is finite and the graph of the affine function obtained by increasing s by $h(x) - s(x)$ is a supporting hyperplane to epigraph h at $(x, h(x))$, where epigraph h is the convex subset of $E \times R$ consisting of all points on or above graph h.

We let $\partial h(x)$ denote the set of subgradients of h at x. This is a weak*-closed convex subset of E^* (which may be empty). The *subdifferential* of h is the following subset of $E \times E^*$:

$$\partial h = \{(x, s) \colon s \in \partial h(x)\} .$$

In [2] Rockefellar applied the methods and results of [1] to the problem of existence of subgradients and showed that a lower semicontinuous proper convex function on a Banach space has a subgradient at a dense set of points in its effective domain (the convex set where it is finite).

Our purpose in this paper is to obtain a relation between the subgradients of a lower semicontinuous proper convex function h and its directional derivatives. A consequence of this will be a formula for the directional derivative of h at a point, in terms of the slopes in the same direction of subgradients at nearby points. This formula will allow us to reduce several questions about lower semicontinuous convex functions on a Banach space to simple questions about convex functions on the real line. In particular it will give us a proof of the maximal monotonicity of the subdifferential of h and of the uniqueness up to an additive constant of h with a given subdifferential.

The result contained in the following lemma and the method of proof were inspired by [1, Theorem 2]. Following the notation of that paper we define for $f \in E^*$, $\|f\| = 1$ and $k > 0$,

$$K(f, k) = \{x \in E\colon \|x\| < kf(x)\} \ .$$

Then $K(f, k)$ is a closed convex cone and, if $k > 1$, has nonempty interior; indeed any x for which $\|x\| < kf(x)$ is in interior $K(f, k)$.

LEMMA 1. *Suppose E is a real Banach space, C is a closed convex subset of E containing a point z, X is a nonempty bounded subset of E and N is a number such that*

$$\sup \{\|z - x\|\colon x \in X\} \leqq N - 1 \ .$$

Suppose $f \in E^*$ and let

$$\delta = \inf f(X) - \sup f(C)$$

$$\theta = \sup f(C) - f(z) \ .$$

Suppose

(1) $$0 < \varepsilon \leqq 1 \quad and \quad \theta < \frac{\delta \varepsilon}{2N} \ .$$

Then there exists $w \in C$ and $g \in E^*$ supporting C at w such that $\|w - z\| \leqq \varepsilon$ and $\sup g(C) < \inf g(X)$.

Proof. From (1) $\delta > 0$, so $f \neq 0$. By dividing f by $\|f\|$ we may (and will) suppose $\|f\| = 1$. (This does not affect (1) since both θ and δ are divided by $\|f\|$.)

Let U be a ball of radius r, $0 < r \leqq 1$, such that if $X' = X + U$ and $\delta' = \inf f(X') - \sup f(C)$ then $\theta < \delta' \varepsilon / 2N$. Let $k = 2N/\delta'$. We notice $k > 1$. Indeed if $x \in X$ then

$$\delta' \leqq \delta \leqq f(x) - f(z) \leqq \|f\| \|x - z\| = \|x - z\| \leqq N - 1 < 2N \ .$$

Using [1, Lemma 1] choose $w \in K(f, k) + z$ so that $K(f, k) + w$ supports C at w. Then $w \in C$ and $w - z \in K(f, k)$ so that

$$\|w - z\| \leqq kf(w - z) \leqq k\theta \leqq \frac{2N}{\delta'} \cdot \frac{\delta'\varepsilon}{2N} = \varepsilon .$$

Now we notice that $X' \subset K(f, k) + w$. Indeed if $x \in X'$ then

$$
\begin{aligned}
\|x - w\| &\leqq \|x - z\| + \|z - w\| \\
&\leqq N + k\theta && (r \leqq 1) + \text{(above calculation)} \\
&< \frac{k\delta'}{2} + \frac{k\delta'\varepsilon}{2N} \\
&< k\delta' && (\text{since } \varepsilon/N < 1) \\
&\leqq kf(x - w) .
\end{aligned}
$$

Since $k > 1$, we have interior $(K(f, k) + w) \neq \phi$ and we can, by [3, 14.2], choose $g \in E^*$ with $\|g\| = 1$ and $\sup g(C) \leqq \inf g(K(f, k) + w)$. Then

$$\sup g(C) \leqq \inf g(K(f, k) + w) \leqq \inf g(X') < \inf g(X) ,$$

the second inequality since $X' \subset K(f, k) + w$ and the third since $\|g\| = 1$ and $r > 0$.

If h is a proper convex function on E, and $y, u \in E, u \neq 0$ and $h(y) < \infty$, we define the derivative of h at y along u to be

$$h'(y; u) = \lim_{\lambda \downarrow 0} [h(y + \lambda u) - h(y)]/\lambda .$$

Since h is convex the limit always exists.

THEOREM 1. *Suppose h is a proper convex lower semicontinuous function on a real Banach space E. Suppose $u \in E, u \neq 0, y \in E$ and $h(y) < \infty$. Suppose $-\infty < d \leqq \infty$ and*

$$(2) \qquad h(y + tu) \geqq h(y) + td \qquad \textit{for all} \quad t \neq 0 .$$

(In particular this is the case if $h'(y; u) = d$.) *If $\varepsilon > 0$ then there exists $v \in E$ and a subgradient s of h at v such that*

$$(3) \qquad \qquad \|v - y\| \leqq \varepsilon$$

$$(4) \qquad \qquad |s(v - y)| \leqq \varepsilon$$

and

$$
\begin{aligned}
&(5.1) && \left\{\begin{array}{ll} |s(u) - d| \leqq \varepsilon & (d < \infty) \\[2mm] s(u) \geqq \dfrac{1}{\varepsilon} & (d = \infty) . \end{array}\right. \\
&(5.2) &&
\end{aligned}
$$

Proof. By translating h and adding a constant we may clearly suppose $y = 0$ and $h(y) = 0$. Then we will be assuming

$(2)'$ $$h(tu) \geqq td \qquad \text{for} \quad t \neq 0$$

and trying to prove

$(3)'$ $$\|v\| \leqq \varepsilon$$

$(4)'$ $$|s(v)| \leqq \varepsilon$$

and either (5.1) or (5.2). Finally we observe that since $u \neq 0$ it is enough to prove the theorem when $\|u\| = 1$. We also assume $\varepsilon < 1$.

We begin by assuming $d = 0$. In the Banach space $E \times R$ with sup norm, let $C = \text{epigraph } h$, $e = (0, 1)$, $\bar{u} = (u, 0)$, $X'' = \{t\bar{u}: -1 \leqq t \leqq 1\}$, $X' = X'' - \varepsilon^2 e/24$ and $X = X'' - \varepsilon e/2$. By $(2)'$ the line $\{t\bar{u}: t \in R\}$ lies under C, so the convex hull C' of C and $\bar{u}$ is disjoint from X'. Since X' is compact convex, and C' is closed we can, by [3, 14.4], choose $f \in (E \times R)^*$ such that $\sup f(C') < \inf f(X')$.

We will verify the hypotheses of Lemma 1 for C, X and f, with $z = (0, 0)$, $N = 2$ and $\varepsilon/2$ instead of ε. Certainly since $z = 0, \varepsilon \leqq 1$ and $\|\bar{u}\| = \|u\| = 1$,

$$\sup\{\|z - x\|: x \in X\} = \sup\{\|x\|: x \in X\} < \sup\left\{\frac{1}{2}, \|\bar{u}\|\right\} = 1 \ .$$

Also if $\delta = \inf f(X) - \sup f(C)$ then

$$\begin{aligned} \delta &\geqq \inf f(X) - \sup f(C') \\ &> \inf f(X) - \inf f(X') \qquad\qquad \text{(by choice of } f) \\ &= \left(\frac{\varepsilon}{2} - \frac{\varepsilon^2}{24}\right)f(-e) > \frac{\varepsilon}{3}f(-e) \ . \end{aligned}$$

Hence

$$\begin{aligned} \theta &= \sup f(C) - f(z) = \sup f(C) \\ &< \inf f(X') \leqq \frac{\varepsilon^2}{24}f(-e) \\ &< \frac{\varepsilon\delta}{8} = \frac{\varepsilon\delta}{4N} \qquad\qquad \text{(by above calculation)} \ . \end{aligned}$$

So by Lemma 1 we can choose $w \in C$ and $g \in (E \times R)^*$ supporting C at w such that $\|w\| = \|w - z\| \leqq \varepsilon/2$ and

(6) $$\sup g(C) < \inf g(X) \ .$$

Let $v \in E$ be the first coordinate of w; then $\|v\| \leqq \|w\| \leqq \varepsilon/2$ giving us $(3)'$. From (6) $g(e) \neq 0$ and so $\{g = 0\}$ is the graph of a linear

function s on E. Since g is continuous the nullspace $\{g = 0\} \cap E$ of s is closed and s is continuous by [3, 5.4]. So $s \in E^*$ and is a subgradient of h at v. Let p be the function on E whose graph is $\{g = g(w)\}$. Then $s = p - p(0)$. Since graph p separates C and X, it lies below $(0, 0)$ but above the points $(\pm u, -\varepsilon/2)$. Thus $|s(u)| = |p(u) - p(0)| \leqq \varepsilon/2$, giving us (5.1). Finally we deduce (4)': since $p(v)$ is the second coordinate of w, $|p(v)| < \|w\| \leqq \varepsilon/2$. Since graph p separates $(0, 0)$ and $(0, -\varepsilon/2)$, we have, $0 \geqq p(0) \geqq -\varepsilon/2$ and therefore

$$|s(v)| = |p(v) - p(0)| \leqq |p(v)| + |p(0)| \leqq \varepsilon .$$

This completes the proof for $d = 0$. Now for arbitrary finite d, choose $\gamma \in E^*$ such that $\gamma(u) = d$ and $\|\gamma\| = d$, (use Hahn-Banach). Let $h' = h - \gamma$. Then

$$h'(tu) = h(tu) - \gamma(tu) = h(tu) - td \geqq 0$$

by $(2)'$. We can apply the theorem for $d = 0$ to h' and choose $v \in E$ and $s' \in E^*$ such that $s'(x - v) \leqq h'(x) - h'(v)$, and

$$\|v\| \leqq \frac{\varepsilon}{(2d + 1)}$$

$$|s'(v)| \leqq \frac{\varepsilon}{2}$$

$$|s'(u)| \leqq \varepsilon .$$

Then setting $s = s' + \gamma$ we have

$$s(x - v) = s'(x - v) + \gamma(x - v) \leqq h'(x) - h'(v) + \gamma(x) - \gamma(v)$$
$$= h(x) - h(v) ,$$

also $\|v\| \leqq \varepsilon$ and

$$|s(v)| = |s'(v) + \gamma(v)| \leqq |s'(v)| + |\gamma(v)|$$
$$\leqq \frac{\varepsilon}{2} + \|\gamma\| \cdot \|v\| \leqq \frac{\varepsilon}{2} + \frac{d\varepsilon}{(2d + 1)} < \varepsilon ,$$

and

$$|s(u) - d| = |s'(u) + \gamma(u) - d| = |s'(u)| \leqq \varepsilon .$$

Finally we assume $d = \infty$. Again we take the Banach space $E \times R$ with sup norm and let $C = $ epigraph h, $e = (0, 1)$, $\bar{u} = (u, 0)$, $X'' = \{te + t\varepsilon\bar{u}/2 : 0 \leqq t \leqq 1\}$, $X' = X'' - \varepsilon^2 e/24$ and $X = X'' - \varepsilon e/2$. By $(2)'$ $h(tu) = \infty$ for $t > 0$, so the convex hull C' of C and X'' is disjoint from X'. Since X' is compact convex, and C' is closed we can, by

[3, 14.4], choose $f \in (E \times R)^*$ such that $\sup f(C') < \inf f(X')$.

Just as in the case $d = 0$, we verify the hypotheses of Lemma 1 for C, X and f, with $z = (0, 0)$, $N = 2$ and $\varepsilon/2$ instead of ε. Certainly

$$\sup \{\| x \|: x \in X\} = \max \left\{ \left\| \frac{-\varepsilon e}{2} \right\|, \left\| \left(1 - \frac{\varepsilon}{2}\right)e + \frac{\varepsilon \bar{u}}{2} \right\| \right\} \leqq 1$$

since $\varepsilon < 1$ and $\|e\| = \|\bar{u}\| = 1$. The computation that $\theta < \delta\varepsilon/4N$ is identical with that for the case $d = 0$. Therefore by Lemma 1 we have $w \in C$ and $g \in (E \times R)^*$ supporting C at w such that $\|w\| \leqq \varepsilon/2$ and

$$(6) \qquad\qquad\qquad \sup g(C) < \inf g(X) .$$

Let $v \in E$ be the first coordinate of w; then $\|v\| \leqq \|w\| \leqq \varepsilon/2$ giving us (3)'. From (6) and the fact that $-\varepsilon e/2 \in X$ and $(0, 0) \in C$, we deduce $g(e) \neq 0$ and so $\{g = 0\}$ is the graph of a linear function s on E. Since g is continuous, the nullspace $\{g = 0\} \cap E$ of s is closed and s is continuous [3, 5.4]. So $s \in E^*$ and is a subgradient of h at v. Let p the function on E whose graph is $\{g = g(w)\}$. Then $s = p - p(0)$. Since graph p separates C and X, graph p lies below $(0, 0)$ but above $(\varepsilon u/2, 1 - \varepsilon/2)$. Thus $p(\varepsilon u/2) - p(0) \geqq 1 - \varepsilon/2$. We deduce

$$s(u) = \frac{2s\left(\dfrac{\varepsilon u}{2}\right)}{\varepsilon} = \frac{2\left(p\left(\dfrac{\varepsilon u}{2}\right) - p(0)\right)}{\varepsilon} > \frac{1}{\varepsilon} ,$$

giving us (5.2). Finally we deduce (4)' by an argument identical to that for the case $d = 0$. This completes the proof of Theorem 1.

If $h'(y; u) = d$ then Theorem 1 gives us a subgradient s at a point close to y whose derivative $s(u)$ along u is arbitrarily close to d. It is natural to ask whether the value of $h'(y; u)$ is determined by the derivatives along u of subgradients of h at points close to y. In case h is finite and continuous one has the following simple formula [7, p. 65]:

$$(7) \qquad\qquad h'(y; u) = \max \{s(u): s \in \partial h(y)\} .$$

One can deduce this formula by noticing that the convex hull of epigraph h and the tangent line through $(y, h(y))$ in direction u with slope $h'(y; u)$ is a convex body with boundary point $(y, h(y))$, hence by [6, p. 72, Prop. 3], has a hyperplane of support at $(y, h(y))$.

If h is only lower semicontinuous, (7) makes no sense since $\partial h(y)$ may be empty. But one can still try to get a formula by using some notation of "approximate subgradient". This general idea and our notation are from Rockafellar [2, 3.3]. (But our $\partial_\varepsilon h$ is not the same

as his.)   For $y \in E$ and $\varepsilon > 0$ we define

$$\partial_\varepsilon h(y) = \{s \in E^*: \exists x \in E \quad \text{with} \quad s \in \partial h(x), \, ||x - y|| < \varepsilon \quad \text{and} \quad |s(x - y)| < \varepsilon\} \, .$$

COROLLARY 1. *If h is a proper convex lower semicontinuous function on a real Banach space E and if $y, u \in E$, $u \neq 0$ and $h(y) < \infty$ then*

(8) $$h'(y; u) = \limsup_{\varepsilon \to 0} \{s(u): s \in \partial_\varepsilon h(y)\} \, .$$

Proof. Theorem 1 tells us that the limit is at least as big as $h'(y; u)$. In case $h'(y; u) = \infty$ we are finished. In case $d = h'(y; u)$ is finite, we get the reverse inequality by choosing any $\varepsilon > 0$. Then we can find $t > 0$ such that

$$h(y + tu) < h(y) + t(d + \varepsilon)$$

and we can choose λ such that $0 < \lambda < t\varepsilon$ and

$$||x - y|| < \lambda \implies h(x) > h(y) - t\varepsilon$$

(by the lower semicontinuity of h). Then if $||x - y|| < \lambda$ and $s \in \partial h(x)$ with $|s(x - y)| < \lambda$ we have

$$\begin{aligned}
h(y + tu) &\geq h(x) + s(y + tu - x) \\
&= h(x) + ts(u) + s(y - x) \\
&> h(y) - t\varepsilon + ts(u) - \lambda \\
&> h(y + tu) - t(d + \varepsilon) - t\varepsilon + ts(u) - \lambda \\
&> h(y + tu) - td + ts(u) - 3t\varepsilon \, .
\end{aligned}$$

Since $t > 0$, $s(u) < d + 3\varepsilon$, and since ε is arbitrary the limit is less than or equal to d.

Several known results follow easily from Corollary 1. In particular two of the results of [2] about lower semicontinuous proper convex functions on Banach spaces are obvious corollaries. The first of these results is that the set of points where such a function has subgradients is dense in its effective domain; the second [2, condition (B)] is that such a function is the supremum of the supporting affine functions determined by its subgradients. To get the second, one must use again the lower semicontinuity of h.

The following corollaries were first announced in [4] but as was pointed out in [5], the proofs given were incomplete. The proofs given in [5, Theorem A and Theorem B] depend upon results about conjugate convex functions.

It is easy to show that the subdifferential of a convex function h is a *monotone* relation i.e.,

$$(x, \alpha) \quad \text{and} \quad (y, \beta) \in \partial h \implies (\alpha - \beta)(x - y) \geqq 0 .$$

COROLLARY 2. [5, Theorem A]. *If h is a lower semicontinuous proper convex function on a Banach space E then ∂h is a maximal monotone relation on $E \times E^*$.*

Proof. We follow the argument of [4, Theorem 4] up to (5.6). The argument first observes that it suffices to show that if $(0, 0) \notin \partial h$ then there is $(x, s) \in \partial h$ such that $s(x) < 0$. It then shows that $(0, 0) \notin \partial h$ implies the existence of $y \in E$, $y \neq 0$ such that $h(y) < \infty$ and $d = h'(y; -y) > 0$. Using (8) we choose $x \in E$ and $s \in \partial h(x)$ such that $|s(x - y)| < d/2$ and $s(-y) > d/2$. Then

$$s(x) = s(x - y) + s(y) < \frac{d}{2} - \frac{d}{2} = 0 .$$

COROLLARY 3. [5, Theorem B]. *If f and h are lower semicontinuous proper convex functions on a Banach space E and if $\partial f = \partial h$ then f and h differ by a constant.*

Proof. If $\partial f = \partial h$ then first of all f and h have the same effective domain. Indeed, by the result of [2] previously quoted, the set where f (or h) has subgradients is dense in its effective domain; since the effective domains are closed (lower semicontinuity), we conclude they are equal. Since $\partial f = \partial h$, $\partial_\varepsilon f(x) = \partial_\varepsilon h(x)$ for all $x \in E$ and $\varepsilon > 0$. Hence by Corollary 1, if $u \neq 0$ then $f'(y; u) = h'(y; u)$ whenever f and h are finite at y. Suppose $x \neq y$ and f and h are finite at x and y. Let $u = x - y$. Then the convex functions $\hat{f}(t) = f(y + tu)$ and $\hat{h}(t) = h(y + tu)$ are finite and lower semicontinuous on $[0, 1]$ and have the same right derivative. Hence

$$\hat{f}(1) - \hat{f}(0) = \int_0^1 \hat{f}^r(t)dt = \int_0^1 \hat{h}^r(t)dt = \hat{h}(1) - \hat{h}(0)$$

and so $f(x) - f(y) = h(x) - h(y)$. (It is not hard to prove the above "Fundamental Theorem of Calculus" for finite lower semicontinuous (hence continuous) convex functions on a closed interval. The superscript r means right derivative).

REMARKS. I am grateful to Professor Rockafeller for his interest in this paper. He supplied the simple proof of Corollary 1 and pointed out the possibility of applying Theorem 1 to get the results of [2].

I am also grateful to the referee for pointing out several places where the results could be expanded or where the style could be improved.

REFERENCES

1. E. Bishop and R. R. Phelps, *The support functionals of a convex set*, Convexity, Proc. Sympos. Pure Math. Vol. 7, Amer. Math. Soc., Providence, R.I., 1962; pp. 27-35.

2. A. Brøndsted and R. T. Rockafellar, *On the subdifferentiability of convex functions*, Proc. Amer. Math. Soc., **16** (1965), 605-611.

3. J. L. Kelly and I. Namioka, *Linear Topological Spaces*, D. Van Nostrand Co. Inc., Princeton, N. J. (1963),

4. R.T. Rockafellar, *Characterization of the subdifferentials of convex functions*, Pacific J. Math., **17** (1966), 497-510.

5. ――――, *On the maximal monotonicity of subdifferential mappings*, Pacific J. Math., **33** (1970), 209-216.

6. N. Bourbaki, *Espaces vectorielles topologiques*, Chs. I-II. Paris: Hermann et Cie, 1953.

7. J.-J. Moreau, *Fonctionelles convex*, mimeographed lecture notes, Collège de France, 1967.

QUEEN'S UNIVERSITY, KINGSTON, ONTARIO