Pergamon

Nonlinear Analysis, Theory, Methods & Applications, Vol. 26, No. 2, pp. 343–350, (1996)

0362-546X/96 $9.50 + .00

0362-546X(94)00262-2

A NONLINEAR DUALITY RESULT EQUIVALENT TO THE CLARKE–LEDYAEV MEAN VALUE INEQUALITY

A. S. LEWIS and D. RALPH

Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ont. N2L 3G1, Canada; and
Department of Mathematics, University of Melbourne, Parkville, Vic. 3052, Australia

(*Received* 18 *February* 1994; *received for publication* 13 *September* 1994)

Key words and phrases: Mean value theorem, Fenchel duality.

1. INTRODUCTION

The main purpose of this paper is to observe the equivalence between the following two results. The first result is the finite-dimensional case of a new mean value theorem due to Clarke and Ledyaev. We use ∂ to denote the Clarke derivative [1].

THEOREM 1 [2, corollary 4.1]. Let X and Y be nonempty, convex, compact sets in $\mathbb{R}^m$, and let Z be the convex hull of $X \cup Y$. Let f be a real function, Lipschitz on a neighbourhood of Z. Then there exists a point z in Z and an element ζ of $\partial f(z)$ with

$$\langle \zeta, y - x \rangle \geq \min_Y f - \max_X f \qquad \text{for all } y \in Y \qquad \text{and} \qquad x \in X.$$

THEOREM 2. Let C be a nonempty, convex, compact set in $\mathbb{R}^m$. Let the functions ϕ, ψ : $\mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be closed, proper and convex, with domains contained in C. Let θ be a real function, Lipschitz on a neighborhood of C. If

$$\phi \geq \theta \geq -\psi \text{ on } C,$$

then there exists a point c in C and an element ξ of $\partial\theta(c)$ with

$$\phi^*(\xi) + \psi^*(-\xi) \leq 0. \tag{1}$$

We begin with a brief discussion highlighting the case where θ is continuously differentiable. Theorem 1 is a powerful generalization of the classical mean value theorem, the latter following easily when X and Y are singletons. Theorem 2 is a nonlinear variant of the fundamental Fenchel duality result. To see this, recall that if

$$\inf_x [\phi(x) + \psi(x)] \geq 0,$$

then, under a regularity condition, Fenchel duality says that there exists a vector u in $\mathbb{R}^m$ such that

$$\phi^*(u) + \psi^*(-u) \leq 0. \tag{2}$$

Theorem 2 gives the additional information that this vector u can be chosen in the range of $\theta'|_C$. Some regularity condition (such as the compactness of C) is clearly needed in general: if we allow $C = \mathbb{R}$, the result can fail (see the example after theorem 7).

Theorem 2 shows that there is an 'affine separator' of ϕ and $-\psi$ (in other words, an affine function lying between the functions ϕ and $-\psi$) which is parallel to the linear approximant to θ at some point c in C. To see this, observe that u is the gradient of an affine separator exactly when, for some constant r

$$-\psi(x) \leq \langle u, x\rangle + r \leq \phi(x), \qquad \text{for all } x,$$

or in other words

$$\sup_x\{-\langle u, x\rangle - \psi(x)\} \leq r \leq \inf_x\{-\langle u, x\rangle + \phi(x)\}.$$

Thus the set of gradients of affine separators is exactly

$$U = \{u \in \mathbb{R}^m | \phi^*(u) + \psi^*(-u) \leq 0\}. \tag{3}$$

To prove theorem 2 it actually suffices to consider functions ϕ and ψ which are continuous on C. This is a consequence of the following simple idea. For a constant $k \geq 0$ we define the *Lipschitz regularization*, $\phi_k : \mathbb{R}^m \to [-\infty, +\infty]$ by

$$\phi_k(x) \stackrel{\text{def}}{=} \inf_y\{\phi(y) + k\|x - y\|\}.$$

The following easy result is standard.

PROPOSITION 3. Suppose that the function $\phi : \mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ is proper and convex, with bounded domain. Then for any $k \geq 0$, the Lipschitz regularization ϕ_k is an everywhere finite convex function with Lipschitz constant k, satisfying $\phi_k \leq \phi$. Suppose furthermore that the set C contains $\operatorname{dom}\phi$, and that the function $\theta : C \to \mathbb{R}$ has Lipschitz constant k and satisfies $\theta \leq \phi$ on C. Then in fact $\theta \leq \phi_k$ on C.

Proof. The function ϕ_k is convex since it is an inf-convolution [3, theorem 5.4], and clearly

$$\phi_k(x) = \inf_y\{\phi(y) + k\|x - y\|\} \leq \phi(x),$$

for all x in $\mathbb{R}^m$. Since ϕ is proper, $\phi_k(x) < +\infty$. On the other hand, since there exists a vector z in $\mathbb{R}^m$ with $\phi(y) \geq \langle z, y\rangle - \beta$ for all y in $\mathbb{R}^m$, [3, corollary 12.1.2], it follows that for all x in $\mathbb{R}^m$

$$\phi_k(x) \geq \inf\{\langle z, y\rangle - \beta + k\|x - y\| \,|\, y \in \operatorname{cl}(\operatorname{dom}\phi)\} > -\infty.$$

Now suppose that two points u and v satisfy $\phi_k(v) < \phi_k(u) - k\|u - v\|$. Then for some w in $\mathbb{R}^m$ we have

$$\phi(w) + k\|v - w\| < \phi_k(u) - k\|u - v\| \leq \phi(w) + k\|u - w\| - k\|u - v\|,$$

contradicting the triangle inequality. Finally, if $\phi_k(x) < \theta(x)$ for some x in C then there exists a point y in $\operatorname{dom}\phi$ with

$$\theta(x) > \phi(y) + k\|x - y\| \geq \theta(y) + k\|x - y\|,$$

contradicting the Lipschitz property of θ. ■

Thus with the assumptions of theorem 2, using this result we can find continuous convex functions ϕ_k and ψ_k with

$$\phi \geq \phi_k \geq \theta \geq -\psi_k \geq -\psi \text{ on } C.$$

Applying theorem 2 to these new functions gives a point c in C and an element ξ of $\partial\theta(c)$ with

$$0 \geq (\phi_k)^*(\xi) + (\psi_k)^*(-\xi) \geq \phi^*(\xi) + \psi^*(-\xi),$$

so the result follows for the original functions.

The proof of theorem 1 in [2] is not all straightforward, involving control–theoretic ideas and a fixed point argument. Theorem 2, unfortunately, does not seem any easier in general. However, in the case $m = 1$ with θ continuously differentiable there is an easy argument. For the purposes of this proof, ∂ will denote the usual convex subdifferential. Given the above comments, we can assume that the functions ϕ and ψ are continuous on the compact interval C, and hence ϕ^* and ψ^* are everywhere finite and continuous.

Classical Fenchel duality shows that the set U given by (3) is a nonempty closed interval. If there is no c in C with (1) holding then without loss of generality, by the intermediate value theorem, we may as well assume that

$$\theta'(z) > \delta \overset{\text{def}}{=} \max U < +\infty, \qquad \text{for all } z \text{ in } C.$$

Define the continuous convex function $\pi : \mathbb{R} \to \mathbb{R}$ by $\pi(u) \overset{\text{def}}{=} \phi^*(u) + \psi^*(-u)$. Since $\delta = \max\{u \in \mathbb{R}^m | \pi(u) \leq 0\}$ it follows that $\pi(\delta) = 0$ and that the right derivative $\pi'_+(\delta) \geq 0$. Hence for some z in $\partial\pi(\delta)$, we have $z \geq 0$.

By the subgradient sum formula there exist z_1 in $\partial\phi^*(\delta)$ and z_2 in $\partial\psi^*(-\delta)$ with $z_1 - z_2 = z \geq 0$. However, then z_1 and z_2 lie in C with

$$\begin{aligned}\delta(z_1 - z_2) &= \phi(z_1) + \phi^*(\delta) + \psi(z_2) + \psi^*(-\delta)\\ &= \phi(z_1) + \psi(z_2) \geq \theta(z_1) - \theta(z_2),\end{aligned}$$

which contradicts the classical mean value theorem if $z_1 > z_2$. On the other hand, if $z_1 = z_2$ then we obtain $\phi(z_1) = \theta(z_1) = -\psi(z_1)$ from the above. Since $\phi \geq \theta \geq -\psi$, it is standard that $\theta'(z_1) \in \partial\phi(z_1)$ and $-\theta'(z_1) \in \partial\psi(z_1)$. However, now

$$\phi^*(\theta'(z_1)) + \psi^*(-\theta'(z_1)) = -\phi(z_1) - \psi(z_1) = 0,$$

contradicting the definition of δ, since $\theta'(z_1) > \delta$.

2. THE EQUIVALENCE OF THEOREM 1 AND THEOREM 2

Proof of theorem 2 from theorem 1. As remarked in the previous section, we can assume that the functions ϕ and ψ are continuous on the set C. Let $\beta \overset{\text{def}}{=} \sup_C \phi$. Then β is finite and we can define a convex, compact, nonempty set

$$Y \overset{\text{def}}{=} \{(y,s) \in \mathbb{R}^{m+1} | \phi(y) \leq s \leq \beta\}.$$

Similarly, let $\alpha \overset{\text{def}}{=} \sup_C \psi$, so

$$X \overset{\text{def}}{=} \{(x,r) \in \mathbb{R}^{m+1} | -\psi(x) \geq r \geq -\alpha\},$$

is a convex, compact, nonempty set.

Now let $f(w,t) \overset{\text{def}}{=} -\theta(w)+t$ for w near C and t in $\mathbb{R}$. Then

$$\inf_Y f = \inf\{-\theta(w)+t \mid \phi(w) \le t \le \beta, w \in C\}$$

$$= \inf\{-\theta(w)+\phi(w) \mid w \in C\} \ge 0,$$

and

$$\sup_X f = \sup\{-\theta(w)+t \mid -\psi(w) \ge t \ge -\alpha, w \in C\}$$

$$= \sup\{-\theta(w)-\psi(w) \mid w \in C\} \le 0.$$

Hence $\inf_Y f - \sup_X f \ge 0$.

By theorem 1 there exist (c,t) in $C \times \mathbb{R}$ and ξ in $\partial\theta(c)$ such that

$$\langle(-\xi,1),(y,s)-(x,r)\rangle \ge 0,$$

for all $x, y \in C$ with $\phi(y) \le s \le \beta$, $\psi(x) \le -r \le \alpha$. Thus

$$[\langle -\xi, y\rangle + \phi(y)] + [\langle \xi, x\rangle + \psi(x)] \ge 0, \qquad \forall x, y \in C,$$

and hence $-\phi^*(\xi) - \psi^*(-\xi) \ge 0$, as required. ■

Proof of theorem 1 from theorem 2. Let X, Y, Z and f be as in theorem 1. Let $\phi \overset{\text{def}}{=} \sup_X f + \delta_X$, where δ_X is the indicator function of X. So $\phi \ge f$ on Z, ϕ is convex, closed and proper. Likewise, $\psi \overset{\text{def}}{=} -\inf_Y f + \delta_Y$ is a convex, closed and proper function, with $f \ge -\psi$ on Z.

By theorem 2, there exists z in Z and ζ in $\partial f(z)$ with

$$\phi^*(\zeta) + \psi^*(-\zeta) \le 0. \tag{3a}$$

Now we have

$$\phi^*(\zeta) = \sup_x\{\langle\zeta, x\rangle - \phi(x)\} = (\sup_{x\in X}\langle\zeta, x\rangle) - \sup_X f,$$

and similarly,

$$\psi^*(-\zeta) = \left(\sup_{y\in Y}\langle -\zeta, y\rangle\right) + \inf_Y f.$$

Substituting these into (3a) yields

$$\inf_Y f - \sup_X f \le -\sup_{y\in Y}\langle -\zeta, y\rangle - \sup_{x\in X}\langle\zeta, x\rangle = \inf_{x\in X, y\in Y}\langle\zeta, y-x\rangle.$$

Theorem 1 follows. ■

3. AN EXTENSION

The requirement in our nonlinear Fenchel result, theorem 2, that the functions ϕ and ψ have domains contained in a compact set appears rather artificial. In this section we consider a version of the result which relies instead on growth conditions on ϕ and ψ. The idea is simple: under reasonable conditions, to find an affine separator for ϕ and ψ it should suffice to separate ϕ and ψ restricted to a large compact set.

We begin by recalling some easy facts about the Lipschitz regularization of a convex function $p\colon \mathbb{R}^m \to \mathbb{R}\cup\{-\infty, +\infty\}$ defined by

$$p_k(u) \stackrel{\text{def}}{=} \inf_v \{ p(v) + k\|u - v\| \}.$$

We denote the closed unit ball in $\mathbb{R}^m$ by B, and δ_{kB} denotes the indicator function of the closed ball with radius k.

Lemma 4. $(p_k)^* = p^* + \delta_{kB}$.

Proof. We have

$$\begin{aligned}(p_k)^*(w) &= \sup_u \{\langle w, u\rangle - p_k(u)\} \\ &= \sup_{u,v} \{\langle w, u\rangle - p(v) - k\|u - v\|\} \\ &= \sup_{z,v} \{\langle w, v + z\rangle - p(v) - k\|z\|\} \\ &= p^*(w) + \delta_{kB}(w),\end{aligned}$$

as required. ■

Lemma 5. If the function p is finite with Lipschitz constant k near the point u then $p_k(u) = p(u)$.

Proof. The convex function $r(v) \stackrel{\text{def}}{=} p(v) + k\|u - v\|$ satisfies $r(v) \geq p(u)$ for v close to u, and $r(u) = p(u)$. Hence r is minimized at u. ■

Lemma 6. Suppose that the convex function p is everywhere finite on $\mathbb{R}^m$, and has bounded level sets. Let α be a real number. Then for all k sufficiently large, $p_k(u) \leq \alpha$ implies that $p_k(u) = p(u)$.

Proof. For β in $\mathbb{R}$, define the level set $L_\beta \stackrel{\text{def}}{=} \{u \mid p(u) \leq \beta\}$, and let $\alpha' \stackrel{\text{def}}{=} \max\{\alpha, p(0)\}$. By [3, theorem 10.4], p is Lipschitz on the bounded level set $L_{\alpha'+2}$, say with Lipschitz constant k_1. Now fix any $k \geq k_1$, and note that $p_k = p$ on $L_{\alpha'+1}$ by lemma 5. We will show that if $p_k(u) \leq \alpha$, then $p_k(u) = p(u)$.

Suppose this fails for some u. Then clearly u does not lie in $L_{\alpha'+1}$, and so $p(u) > \alpha' + 1$. On the other hand, $p(0) \leq \alpha'$, so 0 does lie in $L_{\alpha'+1}$, and hence $p_k(0) = p(0)$. Furthermore, since p is continuous we can choose λ in $(0,1)$ with $\alpha' < p(\lambda u) \leq \alpha' + 1$. Since λu lies in $L_{\alpha'+1}$ we deduce that $p_k(\lambda u) = p(\lambda u)$, whence

$$\alpha' < p(\lambda u) = p_k(\lambda u) \leq (1-\lambda) p_k(0) + \lambda p_k(u) \leq (1-\lambda)\alpha' + \lambda\alpha \leq \alpha',$$

which is a contradiction. ■

We can now give a proof of a variant of theorem 2 involving growth conditions on ϕ and ψ. A convex function ϕ is said to be *cofinite* if it is closed and proper, with *recession function*

$(f0^+)(y) = +\infty$ for all nonzero y, where

$$(f0^+)(y) \stackrel{\text{def}}{=} \lim_{\lambda \to +\infty} \lambda^{-1} f(x + \lambda y),$$

for an arbitrary choice of x in the domain of f. Cofinite convex functions can be characterized as conjugates of everywhere finite convex functions [3, corollary 13.3.1]. They are those proper, closed convex functions which grow faster than linearly. In the following variant of theorem 2 we relax the restriction that the underlying set C be bounded (in particular, we allow $C = \mathbb{R}^m$), at the expense of introducing a constraint qualification and growth conditions.

THEOREM 7. Let C be a nonempty, closed, convex set in $\mathbb{R}^m$. Let the functions ϕ, ψ : $\mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be convex and cofinite, with domains contained in C, and satisfying

$$\text{int}\,(\text{dom}\,\phi) \cap \text{int}\,(\text{dom}\,\psi) \neq \emptyset. \tag{4}$$

Let θ be a real function, locally Lipschitz on a neighbourhood of C. If

$$\phi \geq \theta \geq -\psi \text{ on } C,$$

then there exists a point c in C and an element ξ of $\partial\theta(c)$ with

$$\phi^*(\xi) + \psi^*(-\xi) \leq 0.$$

Proof. By translation we can assume that 0 lies in int(dom ϕ) and int(dom ψ), using (4). Hence ϕ^* and ψ^* have bounded level sets, and are everywhere finite by cofiniteness. If we apply theorem 2 with C replaced by $C \cap kB$, for $k = 1, 2, \ldots$, then for each k we obtain a point c^k in $C \cap kB$ and an element ξ^k of $\partial\theta(c^k)$ with

$$(\phi + \delta_{kB})^*(\xi^k) + (\psi + \delta_{kB})^*(-\xi^k) \leq 0. \tag{5}$$

Now by lemma 4, for all large k we have $(\phi + \delta_{kB})^* = (\phi^*)_k$ and $(\psi + \delta_{kB})^* = (\psi^*)_k$, and furthermore (5) implies that

$$(\phi^*)_k(\xi^k) \leq -(\psi + \delta_{kB})^*(-\xi^k) \leq (\psi + \delta_{kB})(0) = \psi(0).$$

Hence for all k sufficiently large

$$(\phi + \delta_{kB})^*(\xi^k) = (\phi^*)_k(\xi^k) = \phi^*(\xi^k),$$

by lemma 6 with $p \stackrel{\text{def}}{=} \phi^*$ and $\alpha \stackrel{\text{def}}{=} \psi(0)$. Similarly,

$$(\psi + \delta_{kB})^*(-\xi^k) = \psi^*(-\xi^k),$$

for all k sufficiently large, and the result follows by (5). ■

The following example shows that we cannot drop the assumption of cofiniteness in the above result.

Example. Define convex functions ϕ and ψ : $\mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ by $\phi(x) \stackrel{\text{def}}{=} x + \delta_{\mathbb{R}_+}(x)$ and $\psi(x) \stackrel{\text{def}}{=} 2 + x^2/2 + \delta_{\mathbb{R}_+}(x)$, and define a continuously differentiable function $\theta : \mathbb{R} \to \mathbb{R}$ by $\theta(x) \stackrel{\text{def}}{=} x - \exp(-x)$. Then, taking $C \stackrel{\text{def}}{=} \mathbb{R}$, all the assumptions of the theorem are satisfied except that ϕ is not cofinite. Since $\phi^*(y) = +\infty$ unless $y \leq 1$, and $\theta'(c) = 1 + \exp(-c) > 1$ for all c, it follows that

$$\phi^*(\theta'(c)) + \psi^*(-\theta'(c)) = +\infty,$$

for all c.

One of the curious features of theorem 2 is that the usual constraint qualification for Fenchel duality is not required: the existence of the Lipschitz separator θ replaces it. By contrast, our proof of theorem 7 requires the constraint qualification (4). It is unclear to us if this assumption is really required for the result. The following theorem is a partial result in this direction: we can drop the constraint qualification if we assume that the separator θ is *globally* Lipschitz. Comparing theorem 2 and the following result, boundedness of C in the former has been replaced by cofiniteness of ϕ and ψ in the latter.

THEOREM 8. Let C be a nonempty, closed, convex set in $\mathbb{R}^m$. Let the functions $\phi, \psi : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be convex and cofinite with domains contained in C. Let θ be a real function, Lipschitz (globally) on a neighbourhood of C. If

$$\phi \geq \theta \geq -\psi \text{ on } C,$$

then there exists a point c in C and an element ξ of $\partial\theta(c)$ with

$$\phi^*(\xi) + \psi^*(-\xi) \leq 0.$$

Proof. Choose any points x_1 in dom ψ and x_2 in dom ϕ. Let k be the Lipschitz constant for θ, and choose any $r > \|x_1 - x_2\|$. Define a function

$$\phi_0(x) \stackrel{\text{def}}{=} \inf_{\|y-x\|\leq r} \{\phi(y) + k\|x-y\|\},$$

and note that $\phi_0 \leq \phi$. For any point x close enough to x_1, we have that $\|x_2 - x\| < r$, and it follows that

$$\phi_0(x) \leq \phi(x_2) + k\|x - x_2\| < +\infty,$$

so x_1 lies in the interior of dom ϕ_0. Now for any point y in C we have that

$$\phi(y) + k\|x-y\| \geq \theta(y) + k\|x-y\| \geq \theta(x),$$

so $\phi_0 \geq \theta$.

Notice that if we define a function $g(x) \stackrel{\text{def}}{=} k\|x\| + \delta_{rB}(x)$, then clearly g^* is everywhere finite, and since ϕ is cofinite we also know that ϕ^* is everywhere finite. Since ϕ_0 is the infimal convolution of ϕ and g, it is a closed, convex function with $\phi_0^* = \phi^* + g^*$ [3, theorem 16.4], so in fact ϕ_0 is also cofinite.

Similarly, if ψ_0 is the infimal convolution of ψ and g, then it is a closed, convex, cofinite function such that $-\psi \leq -\psi_0 \leq \theta$ on C. As $x_1 + rB \subset \text{dom}\ \psi_0$, the intersection of int(dom ϕ_0) and int(dom ψ_0) is nonempty. Thus we can apply theorem 7 (with ϕ and ψ respectively replaced by ϕ_0 and ψ_0) to deduce the existence of a point c in C and an element ξ of $\partial\theta(c)$ with

$$\phi^*(\xi) + \psi^*(-\xi) \leq \phi_0^*(\xi) + \psi_0^*(-\xi) \leq 0,$$

as required. ∎

COROLLARY 9. Suppose that the function $\theta : \mathbb{R}^m \to \mathbb{R}$ is continuously differentiable and that for some constants $k \geq 0$, $K > 0$ and $1 < p < +\infty$, the growth condition

$$|\theta(x)| \leq k + K\|x\|_p^p/p$$

holds for all x. Then there exists a point $\bar{x}$ satisfying

$$\|\theta'(\bar{x})\|_q \leq K(kq/K)^{1/q}$$

(where $1/p + 1/q = 1$).

Proof. Let $\phi(x) \stackrel{\text{def}}{=} k + K\|x\|_p^p/p \stackrel{\text{def}}{=} \psi(x)$, so that $\phi^*(y) = -k + K\|K^{-1}y\|_q^q/q = \psi^*(y)$. By theorem 7 there exists a point $\bar{x}$ with $\|\theta'(\bar{x})\|_q^q \leq kqK^{q-1}$. ■

For example, if the real function θ is continuously differentiable with $|\theta(x)| \leq 1 + \|x\|^2$ for all x, then there exists a point $\bar{x}$ with $\|\theta'(\bar{x})\| \leq 2$.

REFERENCES

1. CLARKE F. H., *Optimization and Nonsmooth Analysis*. John Wiley and Sons, New York (1983).
2. CLARKE F. H. & LEDYAEV YU. S., Mean value inequalities, *Proc. Am. math. Soc.* **122**, 1075–1083 (1994).
3. ROCKAFELLAR R. T., *Convex Analysis*. Princeton University Press, Princeton, NJ (1972).