Journal of Convex Analysis
Volume 4 (1997), No. 2, 373–379

Geometric Approximation of Proximal Normals

M. L. Radulescu
Department of Mathematics, University of British Columbia,
Vancouver BC V6T 1Z2, Canada.
e-mail: mihaela@math.ubc.ca

F. H. Clarke
Centre de recherches mathématiques, Université de Montréal,
C. P.6128, succ. Centre-ville, Montréal QC H3C 3J7, Canada.
e-mail: clarke@crm.umontreal.ca

Received September 27, 1996
Revised manuscript received May 23, 1997

For $x \in H \setminus S$ and $\delta \geq 0$, the δ-*projection* of x onto S, is the set $\text{proj}_S^\delta(x) := \{s \in S: \|s - x\|^2 \leq d_S(x)^2 + \delta^2\}$. We prove that each vector $x - s$ with $s \in \text{proj}_S^\delta(x)$ can be approximated by some nearby proximal normal. We also give a simple proof (new in the context of an infinite dimensional Hilbert space) of a result due to Rockafellar [17] concerning the approximation of "horizontal" normals to the epigraph of a lower semicontinuous function by "non-horizontal" ones.

Keywords: Nonsmooth analysis, distance function, δ-projection, proximal normal, proximal subdifferential.

1. Introduction

The purpose of this article is to shed new light on a few fundamental results in proximal analysis, in a Hilbert space setting. We are concerned mainly with the following two issues: first, the approximation of "almost"-proximal normal vectors to a closed subset S of a Hilbert space H by nearby exact ones, and second, a generalization of a result by Rockafellar [17] on the approximation of "horizontal" normals to the epigraph of a lower semicontinuous function by "non-horizontal" ones. Our approach, based on the distance function and its differentiability properties, emphasizes once again the particular relevance of these tools to geometric issues in nonsmooth analysis.

The results have been motivated in part by certain applications to the theory of differential inclusion problems involving invariance of trajectories and monotonicity along trajectories (see [10, 16]).

If H is a Hilbert space and S is a closed nonempty subset of H then the *distance function* d_S is defined by

$$d_S(x) := \inf\{\|x - s\|: s \in S\}.$$

We recall that the distance function is globally Lipschitz of rank 1. The *projection* of x

ISSN 0944-6532 / \$ 2.50

onto S is the set of *closest points* to x in S:

$$\mathrm{proj}_S(x) := \{s \in S\colon \|x - s\| = d_S(x)\}.$$

For $\delta \geq 0$, we introduce the *δ-projection* of x onto S, which is the set

$$\mathrm{proj}_S^\delta(x) := \{s \in S\colon \|s - x\|^2 \leq d_S(x)^2 + \delta^2\}.$$

Note that $\mathrm{proj}_S^\delta(x)$ is always nonempty if $\delta > 0$, while $\mathrm{proj}_S^0(x)$ coincides with the possibly empty set $\mathrm{proj}_S(x)$.

A vector $u \in X$ is said to be a *proximal normal* to S at a point s belonging to S if there are $x \notin S$ and $\lambda > 0$ such that

$$u = \lambda(x - s) \quad \text{and} \quad s \in \mathrm{proj}_S(x).$$

The set of all proximal normals to S at s is denoted $N_S^P(s)$ and is referred to as the *proximal normal cone* to S at s. If $x \in \operatorname{int} S$, or if no proximal normals to S exist at x, then by convention, $N_S^P(x) = \{0\}$.

Let $f\colon H \to (-\infty, \infty]$ be a lower semicontinuous function and $x \in \operatorname{dom} f := \{y\colon f(y) < \infty\}$. An element $\zeta \in H$ is said to be a *proximal subgradient* of f at x provided that

$$(\zeta, -1) \in N_{\operatorname{epi} f}^P(x, f(x))\,.$$

The set of proximal subgradients of f at x, denoted $\partial_P f(x)$, and called the *proximal subdifferential* of f at x can be empty; however, it is nonempty on a dense subset of $\operatorname{dom} f$. The proximal normal cone can be given an analytical expression via the indicator function of the set S:

$$N_S^P(s) = \partial_P \psi_S(s) \quad \forall s \in S. \tag{1.1}$$

The following particular sum rule will also be invoked.

Proposition 1.1. *Let $g\colon H \to \mathbb{R}$ be of class $\mathbb{C}^2$ on an open set Ω and $x_0 \in \Omega$ be a local minimum for the function $f + g$. Then $-g'(x_0) \in \partial_P f(x_0)$.*

We will require the following result of [11] regarding proximal subdifferentiability of the distance function and the existence of nearest points (for related work in Banach space settings see [1, 2, 13, 14]).

Theorem 1.2. *Suppose $x \notin S$ and $\zeta \in \partial_P d_S(x)$. Then*

(a) $\|\zeta\| = 1$, $\mathrm{proj}_S(x)$ *is a singleton* $\{s\}$ *and*

$$\partial_P d_S(x) = \left\{\frac{x - s}{\|x - s\|}\right\} = \{\zeta\} \subseteq N_S^P(s);$$

(b) *any minimizing sequence for the infimum defining $d_S(x)$ converges in norm to s;*

(c) *if $x_n \to x$ and $\zeta_n \in \partial_P d_S(x_n)$ then $\zeta_n \to \zeta$.*

The closed *t-outer approximation* of S ($t \geq 0$) is defined by

$$S(t) := \{x \in H\colon d_S(x) \leq t\}.$$

In [12], the authors gave the following characterization of the subgradient of the distance function.

Theorem 1.3. *Suppose that* $d_S(u) = t > 0$. *Then*

$$\partial_P d_S(u) = N^P_{S(t)}(u) \cap \{\zeta \in H : \|\zeta\| = 1\},$$

where emptiness is not precluded.

More about the geometry of the t-outer approximation was developed in connection with proximally smooth sets and the Complementary Normal Formula (see [12, 8]).
We will make reference to the following consequence of the Mean Value Inequality of [7].

Theorem 1.4. *Let* x, $y \in H$. *Then for all* $r < d_S(y) - d_S(x)$ *and* $\varepsilon > 0$, *there exist* $z \in [x, y] + \varepsilon B$ *and* $\zeta \in \partial_P d_S(z)$ *such that*

$$r < \langle \zeta, y - x \rangle.$$

2. Geometric Approximations

The converse of Theorem 1.2 is not necessary true; i.e. even when $\text{proj}_S(x)$ is a singleton we can have $\partial_P d_S(x) = \emptyset$ (for an example see [11], Remark 4.13). However, a partial converse still holds; it was inplicitly stated in [11]. Here is the explicit formulation.

Proposition 2.1. *Let* $x \notin S$ *and* $s \in \text{proj}_S(x)$. *Then for all* $t \in (0, 1)$,

$$\partial_P d_S(s + t(x - s)) = \left\{ \frac{x - s}{\|x - s\|} \right\}.$$

Proof. We may suppose without loss of generality that $\|x - s\| = 1$. Fix $t \in (0, 1)$ and let $u := s + t(x - s)$. Then $x - u \in N^P_{S(t)}(u)$. (Indeed: we have $d_{S(t)}(x) \leq \|x - u\| = 1 - t$. Suppose that the inequality is strict. Then there exists $x_t \in S(t)$ such that $\|x - x_t\| < 1 - t$ which implies that $d_S(x) \leq d_S(x_t) + \|x - x_t\| < t + (1 - t)$; contradiction.) Now we apply Theorem 1.3 to deduce that $\partial_P d_S(u) \neq \emptyset$. Then by Theorem 1.2 and since $s \in \text{proj}_S(u)$, $\partial_P d_S(u) = \{(u - s)/\|u - s\|\} = \{(x - s)/\|x - s\|\}$ as required. □

We note the crucial role played by the set $S(t)$, which however does not appear explicitly in the above statement.
The following result asserts that approximate projection directions can be estimated by nearby exact ones.

Theorem 2.2. *Let* $d_S(x) > 0$. *Then for any* $\delta \in [0, d_S(x))$ *and* $\varepsilon \in (0, d_S(x) - \delta)$, *for any* $s \in \text{proj}^\delta_S(x)$, *there exist* $\overline{x} \notin S$, $\|\overline{x} - x\| < \delta + 2\varepsilon$ *and* $\zeta \in \partial_P d_S(\overline{x})$ *such that*

$$\left\| \frac{x - s}{\|x - s\|} - \zeta \right\| < \frac{4(\delta + \varepsilon)}{d_S(x)}. \tag{2.1}$$

Further, $\text{proj}_S(\overline{x}) = \{\overline{s}\}$ *for some* $\overline{s} \in S$ *satisfying* $\|s - \overline{s}\| < \delta + \varepsilon$.

Proof. We may suppose $\delta > 0$, since for $\delta = 0$ the conclusion follows from Proposition 2.1. Consider the function $f(y) := \|y - x\|^2 + \psi_S(y)$ where ψ_S is the indicator function of the set S. By hypothesis, we have

$$f(s) < \inf_{y \in S} f(y) + (\delta + \varepsilon)^2.$$

We apply the Borwein-Preiss Variational Principle with $\lambda = \delta + \varepsilon$ (see Theorems 2.6 and 5.2 in [3]) to deduce the existence of points $\bar{s} \in S$ and $\bar{z} \in H$ such that

$$\|\bar{s} - s\| < \delta + \varepsilon, \quad \|\bar{z} - \bar{s}\| < \delta + \varepsilon$$

and such that the following function of y attains a minimum over S at $y = \bar{s}$:

$$\|y - x\|^2 + \|y - \bar{z}\|^2.$$

It follows from Proposition 1.1 and formula (1.1) that $x' - \bar{s} \in N_S^P(\bar{s})$, where

$$x' := x + (\bar{z} - \bar{s}).$$

Because $\|\bar{z} - \bar{s}\| < \delta + \varepsilon < d_S(x)$, we have $x' \notin S$. We now apply Proposition 2.1 to deduce the existence of $\bar{x} \notin S$, $\|\bar{x} - x'\| < \varepsilon$ and $\zeta \in \partial_P d_S(\bar{x})$ such that

$$\zeta = \frac{x' - \bar{s}}{\|x' - \bar{s}\|},$$

where $\operatorname{proj}_S(\bar{x}) = \{\bar{s}\}$. Also,

$$\|\bar{x} - x\| \leq \|\bar{x} - x'\| + \|x' - x\| < \delta + 2\varepsilon,$$

as required. It remains to verify (2.1), which will follow if we prove that the quantity

$$\left\| \frac{x - s}{\|x - s\|} - \frac{x' - \bar{s}}{\|x' - \bar{s}\|} \right\|,$$

is bounded above by $4(\delta + \varepsilon)/d_S(x)$. But the quantity in question can also be written as

$$\begin{aligned}
\left\| \frac{x - s}{\|x - s\|} - \frac{x' - \bar{s}}{\|x' - \bar{s}\|} \right\| &= \frac{\left\| (x' - \bar{s}) - (x - s) + \left(\frac{\|x - s\|}{\|x' - \bar{s}\|} - 1 \right) (x' - \bar{s}) \right\|}{\|x - s\|} \\
&\leq \frac{\|x' - x\| + \|\bar{s} - s\| + \left| \|x - s\| - \|x' - \bar{s}\| \right|}{d_S(x)} \\
&\leq 2 \frac{\|x' - x\| + \|\bar{s} - s\|}{d_S(x)} \\
&= 2 \frac{\|\bar{z} - \bar{s}\| + \|\bar{s} - s\|}{d_S(x)} \\
&< \frac{4(\delta + \varepsilon)}{d_S(x)}.
\end{aligned}$$

□

As a consequence we obtain the Proximal Density Theorem (see [12]; see also [15, 1] for a Banach space version).

Corollary 2.3. *The function d_S has a nonempty subdifferential on a dense subset of $H \setminus S$, so $\operatorname{proj}_S(x) \neq \emptyset$ for all x in a dense subset of $H \setminus S$.*

We remark that Theorem 2.2 is used in [16] to extend the "proximal aiming" technique of [10] to infinite dimensions.

We now address the issue of "horizontal" normals to epigraphs of functions, an important point in developing proximal calculus. By definition, to any proximal subgradient of a lower semicontinuous function at a point $y \in \operatorname{dom} f$ there corresponds a proximal normal to the epigraph of f at $(y, f(y))$. But there are proximal normals to $\operatorname{epi} f$ which do not necessarily correspond to some subgradient of f. These are the so-called "horizontal" normals, vectors of the form $(y^*, 0)$. (Locally Lipschitz functions do not exhibit such anomalies.) However, one can find a nearby subgradient of f defining a nearby "nonhorizontal" normal.

Theorem 2.4. *Let $f : H \to (-\infty, +\infty]$ be a lower semicontinuous function, $y \in \operatorname{dom} f$ and $(y^*, 0) \in N^P_{\operatorname{epi} f}(y, f(y))$ with $y^* \neq 0$. Then for any $\varepsilon > 0$ there exist $\overline{x} \in (y + \varepsilon B) \cap \operatorname{dom} f$ with $|f(\overline{x}) - f(y)| < \varepsilon$, $\lambda \in (0, \varepsilon)$, and $\zeta \in y^* + \varepsilon B$ such that $(\zeta, -\lambda) \in N^P_{\operatorname{epi} f}(\overline{x}, f(\overline{x}))$.*

Proof. Let us set $\varphi := f(y)$, $S := \operatorname{epi} f$ and without loss of generality assume that $\|y^*\| = 1$. Then there is a point $(x, \varphi) \notin S$ having closest point (y, φ) in S, and such that $(x - y)/\|x - y\| = y^*$. Also, in light of Proposition 2.1, there is no loss of generality in supposing that $(y^*, 0)$ belongs to $\partial_P d_S(x, \varphi)$.

Note that for any (x', φ') we have

$$d_S(x', \varphi') \leq d_S(x', \varphi' - t) \quad \forall t > 0,$$

as a consequence of the nature of an epigraph. Suppose that it is possible to find (x', φ') arbitrarily close to (x, φ) and $t > 0$ arbitrarily small so that strict inequality holds in this relation. Then Theorem 1.4 produces an element $(\zeta, -\lambda)$ of $\partial_P d_S(\overline{x}, \overline{\varphi})$, where $(\overline{x}, \overline{\varphi})$ is as close to (x, φ) as we wish, such that

$$0 < \langle (\zeta, -\lambda), (x', \varphi' - t) - (x', \varphi') \rangle.$$

Consequently, $\lambda > 0$. Also, by Theorem 1.2.c, such $(\zeta, -\lambda)$ necessarily converges to $(y^*, 0)$ so the theorem follows.

The only situation to be dealt with, then, is that in which there exists $\delta > 0$ such that

$$\|x' - x\| < \delta, \quad |\varphi' - \varphi| < \delta \Longrightarrow d_S(x', \varphi' - t) = d_S(x', \varphi') \quad \forall t \in [0, \delta]. \tag{2.2}$$

Choose $x_i \to x$ and $t_i \in (\delta/2, \delta)$ so that $(x_i, \varphi - t_i) \notin \operatorname{epi} f = S$ (we can do so by semicontinuity of f) and for certain $y_i \in H$ and $\delta_i \geq 0$,

$$d_S(x_i, \varphi - t_i) = \|(x_i, \varphi - t_i) - (y_i, f(y_i) + \delta_i)\|$$

(by the Proximal Density Theorem). Then in view of (2.2), for large i

$$d_S(x_i, \varphi) = \|(x_i, \varphi) - (y_i, f(y_i) + \delta_i + t_i)\|.$$

Now, since d_S is continuous and $x_i \to x$, we have

$$\|(x, \varphi) - (y_i, f(y_i) + \delta_i + t_i)\| \to d_S(x, \varphi)$$

which means that $f(y_i) + \delta_i + t_i \in S$ is a minimizing sequence for the infimum defining $d_S(x, \varphi)$. Thus

$$(y_i, f(y_i) + \delta_i + t_i) \to (y, \varphi) = (y, f(y)),$$

by Theorem 1.2. But the limit of $f(y_i) + \delta_i + t_i$ must be at least $f(y) + \delta/2$, since $t_i \geq \delta/2$ and f is lower semicontinuous. This contradiction completes the proof. □

Remark 2.5. Rockafellar, in his paper [17], gave an optimization-based proof in the case of $\mathbb{R}^n$. Borwein and Strojwas [5] built upon Rockafellar's idea and proved the result in the framework of a reflexive Banach space with Kadec and Fréchet differentiable norm. The proof is very technical, due not only to the specific geometry of the space (less rich than the geometry of a Hilbert space) but also due to their "constructive" approach. Our approach, based on the differentiability properties of the distance function and on the Mean Value Inequality, is shorter, less technical, but limited to the Hilbert space for the moment.

References

[1] J. M. Borwein, S. Fitzpatrick: Existence of nearest points in Banach spaces, Can. J. Math. 41(4) (1989) 702–720.

[2] J. M. Borwein, J. R. Giles: The proximal normal formula in Banach space, Trans. Amer. Math. Soc. 302(1) (1987) 371–381.

[3] J. M. Borwein, D. Preiss: A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, Trans. Amer. Math. Soc. 303 (1987) 517–527.

[4] J. M. Borwein, H. M. Strojwas: Proximal analysis and boundaries of closed sets in Banach space, Part I: Theory, Can. J. Math. 38 (1986) 431–452.

[5] J. M. Borwein, H. M. Strojwas: Proximal analysis and boundaries of closed sets in Banach space, Part II: Applications, Can. J. Math. 39 (1987) 428–472.

[6] F. H. Clarke: Optimization and nonsmooth analysis, In: Classics in Applied Mathematics 5, SIAM, Philadelphia, 1990. (Originally published by Wiley Interscience, New York, 1983)

[7] F. H. Clarke, Yu. S. Ledyaev: Mean value inequalities in Hilbert space, Trans. Amer. Math. Soc. 344 (1994) 307–324.

[8] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern: Complements, approximations, smoothings and invariance properties. (to appear)

[9] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, P. R. Wolenski: Introduction to Nonsmooth Analysis. (textbook in preparation)

[10] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, P. R. Wolenski: Qualitative properties of trajectories of control systems: a survey, J. Dynamical Control Systems 1 (1995) 1–48.

[11] F. H. Clarke, Yu. S. Ledyaev, P. R. Wolenski: Proximal analysis and minimization principles, J. Math. Anal. Appl. 196 (1995) 722–735.

[12] F. H. Clarke, R. J. Stern, P. R. Wolenski: Proximal smoothness and the lower-C^2 property, J. Convex Anal. 2 (1995) 117–144.

[13] I. Ekeland, G. Lebourg: Generic Fréchet-differentiability and perturbed optimization problems in Banach spaces, Trans. Amer. Math. Soc. 222(2) (1976) 193–216.

[14] A. Ioffe: Approximate subdifferentials and applications II, Mathematika 33 (1986) 111–128.

[15] K. S. Lau: Almost Chebychev subsets in reflexive Banach spaces, Indiana Univ. Math. J. 2 (1978) 791–795.

[16] M. L. Radulescu, F. H. Clarke, Yu. S. Ledyaev: Approximate invariance and differential inclusions in Hilbert spaces. (in preparation)

[17] R. T. Rockafellar: Proximal subgradients, marginal values, and augmented Lagrangians in nonconvex optimization, Math. Oper. Res. 6(3) (1981) 424–436.

HIER :

Leere Seite
380