Revista de la
Unión Matemática Argentina
Volumen 31, 1984.

A NOTE ON THE EXTENSION OF LIPSCHITZ FUNCTIONS

Telma Caputti

1. INTRODUCTION.

In many areas including optimization problems as well as some important questions of analysis, we have to deal with functions F satisfying a Lipschitz property only on a subset S of the whole space E. It is important to know whether F can be extended to E preserving such a property, that is whether there exists a function F_S, defined and possessing a Lipschitz property on all of E, which coincides with F on S. For such a problem, an explicit formula for the extension was given forty - five years ago by E.J.McShane [1] but one can propose here an alternative extension obtained by performing the infimal convolution of two functions associated with the data of the problem. Although conceptually identical to McShane's procedure, the extension by infimal convolution is more suitable for minimization problems. The difference will also appear to be relevant when comparing generalized gradients of the respective functions.

The first section is introductory; the second section deals with the definition and basic properties of the space of Lipschitz functions on a subset. In Section III we introduce the extension process. Section IV is devoted to comparison results between the generalized gradient of the extended function and that of the initial function. In view of applications, we consider in Section V problems dealing with optimization of the extended function. In particular, it will be proved that the search for a global or local minima of F on S is equivalent to the same problem on E with the extension as objective function.

2. LIPSCHITZ FUNCTIONS.

Let E be a real Banach space and let $\|.\|$ denote the norm of E.

Given a nonempty subset S of E, F: E $\rightarrow \overline{R}$ (the extended reals) is said to be Lipschitz on S with Lipschitz constant $r \geqslant 0$ if F is finite on S and if

$$|F(x) - F(y)| \leqslant r \|x-y\| \quad \text{for all} \quad x,y \text{ in } S \tag{2.1}$$

The class of all such functions is denoted by $L_r^{ip}(S)$. The class of all $L_r^{ip}(s)$ for $r \geqslant 0$ is the class of Lipschitz functions on S and is denoted by $L^{ip}(S)$. It is evident that $F \in L^{ip}(S)$ only in the case where

$$|||F||| = \sup\{ \frac{|F(x)-F(y)|}{\|x-y\|} \; ; \; y,x \text{ in } S \; ; \; x \neq y\} < \infty \tag{2.2}$$

$|||F|||$ is the least number r such that (1.1) holds for F.
Suppose that $x \in S$ and define

$$|||F|||_{\overline{x}} = |F(\overline{x})| + |||F||| \quad \text{for all} \quad F \in L^{ip}(S).$$

Then $(L^{ip}(S), |||.|||_{\overline{x}})$ is a Banach space [2].
Since only the values of F on S are relevant for our purpose, we will make a constant use of F defined on E by

$$\overline{F}(x) = F(x) \quad \text{if} \quad x \in S \quad ; \quad +\infty \quad \text{if not} \tag{2.3}$$

In particular, the Lipschitz property of F on S may be expressed in terms of the infimal convolution $F_{S,r}$ which appears as the result of a sort of regularization as follows:
Let F be non identically $(+\infty)$ or $(-\infty)$ on S; then $F \in L_r^{ip}(S)$ if and only if

$$\overline{F} \,\square\, r\,\|.\| = F \quad \text{on} \quad S \tag{2.4}$$

which is, furthermore, equivalent to

$$\overline{F} \,\square\, r\,\|.\| \geqslant F \quad \text{on} \quad S \tag{2.5}$$

where the symbol $\square$ denotes infimal convolution defined by: Let g and h be two functions from E into $\overline{R}$, the infimal convolution of g and h is a function, denoted by $g \,\square\, h$, which assigns to $x \in E$ the value

$$\inf_{u \in E} \{g(u) + h(x-u)\} .$$

The general properties of this binary operation, particularly those related to convex analysis are developed in [3].

3. EXTENSION OF THE RANGE OF A LIPSCHITZ FUNCTION.

Let S be a nonempty subset of E and let $F \in L_r^{ip}(S)$. In 1934, McShane showed that such a function F could be extended to the whole space E by preserving a Lipschitz condition. Actually, his procedure yielded an explicit formula for the extension $F^{S,r}$ which was

$$F^{S,r}(x) = \sup_{u \in S} \{F(u) - r \|x-u\|\} \qquad (3.1)$$

$F^{S,r}$ turns out to be Lipschitz on E with r as Lipschitz constant and coincides with F on S.

We define another extension which is conceptually related to McShane's one [1]. The definition of the extended function $F_{S,r}$ comes naturally from paragraph 2 as

$$F_{S,r} = \overline{F} \,\square\, r \|.\| \qquad (3.2)$$

In more explicit way

$$F_{S,r}(x) = \inf_{u \in S} \{F(u) + r \|x-u\|\} \text{ for all x in E .}$$

Clearly if $F \in L_r^{ip}(S)$ then $F_{S,r} \in L_r^{ip}(S)$ and coincides with F on S.

4. THE GENERALIZED GRADIENT OF THE EXTENDED FUNCTION.

Given a function F Lipschitz in a neighborhood of $x_o \in E$, the generalized gradient of F at x_o in Clarke's sense [5] is a subset of E* (topological dual space of E) denoted by $\partial F(x_o)$ and defined as follows:

$$\partial F(x_o) = \{x^* \in E^*: \langle x^*,d\rangle \leqslant F^\circ(x_o;d) \text{ for all } d \in E\} \qquad (4.1)$$

where

$$F^\circ(x_o;d) = \operatorname*{Lim\,sup}_{\substack{x \to x_o \\ \lambda \to 0}} \frac{F(x+\lambda d) - F(x)}{\lambda} \qquad (4.2)$$

The definition of the generalized gradient for an arbitrary function requires some preliminary definitions. Let E be a real Banach space, let A be a subset of E and let $u_o \in cl(A)$ (closure of A).

DEFINITION 4.3. δ is tangent direction to A at u_o if and only if for every sequence $\{u_n\} \subset A$ converging to u_o and for every sequence $\{\lambda_n\} \subset R_+^*$ converging to 0, there exists a sequence $\{\delta_n\}$ converging to δ such that $u_n + \lambda_n \delta_n \in A$ for all n.

The cone of all tangent directions to A at u_o is the *tangent cone to* A at u_o and will be denoted by $T_A(u_o)$. Its polar cone, i.e. the set of n in E* such that $\langle n,\delta\rangle \leqslant 0$ for all $\delta \in T_A(u_o)$ is called the *normal cone* to A at u_o and will be denoted by $N_A(u_o)$.

Let $F: E \to \overline{R}$ be finite at x_o and Lipschitz around x_o. Starting from the geometric concept of tangent cone, the generalized directional derivative of F at x_o is defined by

$$d \longrightarrow F^{\circ}(x_o;d) = \inf \{\mu \in R: (d,\mu) \in T_{epiF}(x_o,F(x_o))\} \qquad (4.4)$$

The relationship with the normal cone is given as follows:

$$\partial F(x_o) = \{x^* \in E^*: (x^*,-1) \in N_{epiF}(x_o,F(x_o))\} \qquad (4.5)$$

For the indicator function of a subset S

$$\delta(x/S) = 0 \text{ if } x \in S \quad ; \quad \delta(x/S) = +\infty \text{ if } x \notin S$$

one has $\delta(x_o/S) = N(S;x_o)$. For more details on what has been recalled above, see [3].

Concerning the generalized gradients of $\overline{F}$ and $F_{S,r}$ such as defined in the previous paragraph 3, we have a general comparison result:

THEOREM 4.6. *Let* x_o *in* S. *Then,*

a) *for all* $r \geq |||F|||$, $\partial\overline{F}(x_o) \subset \partial F_{S,r}(x_o) + N(S;x_o)$.

b) *for all* $r \geq |||F|||$, $\partial F_{S,r}(x_o) \subset \partial\overline{F}(x_o) \cap rB^*$ *where* B^* *denotes the closed unit ball in* E^**:*

Proof. a) Since $F_{S,r}$ coincides with F on S, we have that $\overline{F} = F_{S,r} + \delta(./S)$. Then the announced result follows from the calculus rule giving an estimate of the generalized gradient of the sum of two functions [6, Theorem 2].

b) $F_{S,r}$ is Lipschitz with constant r, therefore

$$F^{\circ}_{S,r}(x_o;d) \leq r\|d\| \text{ for all } d \text{ and } \partial F_{S,r}(x_o) \subset rB^*$$

If x_o is in int(S), $F_{S,r} = \overline{F} = F$ in a neighborhood of x_o; thus

$$\partial F_{S,r}(x_o) = \partial\overline{F}(x_o) = \partial F(x_o)$$

Let now, x_o in $S \cap bd(S)$; we have $F_{S,r}(x_o) = \overline{F}(x_o) = F(x_o)$; the inclusion

$$\partial F_{S,r}(x_o) \subset \partial\overline{F}(x_o)$$

is then equivalent to the following one

$$T_{epi\overline{F}}(x_o,F(x_o)) \subset T_{epiF_{S,r}}(x_o,F(x_o)) \qquad (4.7)$$

Let (d,μ) in $T_{epi\overline{F}}(x_o,F(x_o))$. We consider a sequence $\{x_n\}$ converging to x_o and a sequence $\{\lambda_n\} \subset R^*_+$ converging to 0. With $\{x_n\}$ and $\{\lambda_n\}$ we associate a sequence $\{\overline{x}_n\} \subset S$ such that $\overline{x} \in M(\lambda_n^2,x_n)$ for all n where M(.,.) is given as in Theorem 5.1.b).

Since $r > |||F|||$, $\{\bar{x}_n\}$ converges to x_o, therefore the sequence $\{(\bar{x}_n, F(\bar{x}_n))\}$ converges to $(x_o, F(x_o))$ in epi $\bar{F}$. Since $(d,\mu) \in T_{\text{epi}\bar{F}}(x_o, F(x_o))$, there exists a sequence $\{(d_n, \mu_n)\}$ converging to (d,μ) such that $\bar{x}_n + \lambda_n d_n \in S$ and

$$F(\bar{x}_n + \lambda_n d_n) \leq F(\bar{x}_n) + \lambda_n \mu_n \text{ for all } n.$$

Due to the Lipschitz property of $F_{S,r}$, we get that

$$F_{S,r}(x_n + \lambda_n d_n) \leq F(\bar{x}_n) + r \, \|x_n - \bar{x}_n\| + \lambda_n \mu_n$$

Since $\bar{x}_n \in M(\lambda_n^2, x_n)$, we have that

$F_{S,r}(x_n + \lambda_n d_n) \leq F_{S,r}(x_n) + \lambda_n(\mu_n + \lambda_n)$. Hence, since $(d_n, \lambda_n + \mu_n) \to (d,\mu), (d,\mu) \in T_{\text{epi}F_{S,r}}(x_o, F(x_o))$ and the inclusion (4.7) is proved. (q.e.d.)

5. OPTIMIZATION OF LIPSCHITZ FUNCTIONS.

Given S a nonempty subset of E and $F \in L_r^{ip}$ we consider the problem of minimizing (at least locally) F on S: (P) minimize F on S.

A device for converting the constrained optimization problem (P) into an unconstrained one is to consider

(P*) minimize $\bar{F}$ on E.

Of course, x_o is a local minimum of F on S if and only if x_o is a local minimum of F on E.

Similar properties hold for the extended function $F_{S,r}$ with the advantage that $F_{S,r}$ is finite and Lipschitz over all E.

THEOREM 5.1. *Let* S *be closed in* E.

a) x_o *is a global minimum of* F *on* S *if and only if* x_o *is a global minimum of* $F_{S,r}$ *on* E $(r > 0)$.

b) x_o *is a local minimum of* F *on* S *if and only if* x_o *is a local minimum of* $F_{S,r}$ *on* E *whenever* $r > |||F|||$.

Proof. a) Let x_o in S such that $F(u) \geq F(x_o)$ for all u in S. Clearly, $F_{S,r}(x) = \inf_{u \in S} \{F(u) + r \, \|x-u\|\} = F(x_o)$ for all x in E.

Conversely, let x_o be a global minimum of $F_{S,r}$ on E. The only thing to prove is that x_o necessarily belongs to S. For that, we suppose that $d_S(x_o) = a > 0$ (distance from x_o to S).

Let $\bar{x}$ in S be such that

$$F_{S,r}(x_o) > F(\bar{x}) + r\,\|\bar{x}-x_o\| - \frac{ra}{2} \qquad (5.2)$$

Since $F_{S,r}$ agrees with F on S and $\bar{x} \in S$, we have that

$$F(\bar{x}) \geqslant F_{S,r}(x_o) \quad \text{and} \quad \|\bar{x}-x_o\| \geqslant a$$

That is inconsistent with inequality (5.2); hence $a=0$ and since S is closed $x_o \in S$.

b) Let x_o in S be a local minimum of F on S. So, there exists $\rho > 0$ such that $F(u) \geqslant F(x_o)$ whenever $u \in S$ and $\|u-x_o\| \leqslant \rho$.

There exists $\rho_o > 0$ and $\varepsilon_o > 0$ such that

$\|x-x_o\| \leqslant \rho_o$; $\varepsilon \leqslant \varepsilon_o \Rightarrow \|\bar{x}-x_o\| \leqslant \rho$ for all $\bar{x}$ in $M(\varepsilon,x)$, where

$$M(\varepsilon,x) = \{u \in S/F(u) + r\,\|x-u\| \leqslant F_{S,r}(x) + \varepsilon\}$$

with $\varepsilon \geqslant 0$ and x in E and $M(x) = M(0,x)$.

Clearly, $M(\varepsilon,x)$ is nonempty for all x in E and all $\varepsilon > 0$. Furthermore, if $x \in S$, $M(x)$ contains x and is reduced to $\{x\}$ whenever $r > |||F|||$.

Generally speaking, computing $F_{S,r}(x)$ gives rise to an abstract optimization problem. It is important to know the behaviour of the set of solutions or of approximate solutions.

Consequently, $F_{S,r}(x) \geqslant F(x_o)$ whenever $\|x-x_o\| \leqslant \rho_o$. Conversely, let us prove that x_o local minimum of $F_{S,r}$ on E is in S. Then exists $\rho > 0$ such that $F_{S,r}(x) \geqslant F_{S,r}(x_o)$ if $\|x-x_o\| \leqslant \rho$. Let us suppose that $d_S(x_o) = a > 0$; we set $\varepsilon < r/2 \min\{\rho,a\}$ and we choose $\bar{x}$ in S satisfying $F_{S,r}(x_o) > F(\bar{x}) + r\|\bar{x}-x_o\| - \varepsilon$.

Let $\theta = 1/2 \min\{\rho,a\}$ and $x^* = x_o + \theta \dfrac{\bar{x} - x_o}{\|\bar{x} - x_o\|}$. We have that

$F_{S,r}(x^*) > F(\bar{x}) + r\,\|\bar{x}-x_o\| - \varepsilon$. Since $x^* \notin S$ and $\|\bar{x}-x_o\| = \|\bar{x}-x^*\| + \theta$ we deduce from (5.2) that $r\theta < \varepsilon$; hence the contradiction from the choice of ε. (q.e.d.)

REMARK. Let $f: O \longrightarrow R$ be a function defined on an open subset O of R^n and Lipschitz in a neighborhood of x_o. Let $B(x_o,\varepsilon)$ be a closed ball around x_o of radius included in O.

We denote by r the Lipschitz constant of f on $B(x_o,\varepsilon)$ and we set

$f^* = f + \delta_{B(x_o,\varepsilon)}$ where $\delta_{B(x_o,\varepsilon)}$ is the indicator function of $B(x_o,\varepsilon)$ defined by

$$\delta_{B(x_o,\varepsilon)}(x) = \begin{cases} 0 & \text{if } x \text{ is in } B(x_o,\varepsilon) \\ +\infty & \text{otherwise.} \end{cases}$$

Now, we perform the infimal convolution of f^* and the function $r\|.\|$, that is

$$\tilde{f}(x) = \inf \{f^*(x_1) + r\,\|x_2\| \ : \ x_1 + x_2 = x\} \ .$$

It is easy to see that by performing this operation we produce a function $\tilde{f}$ such that

i) $\tilde{f}$ is Lipschitz on the whole space with Lipschitz constant r.

ii) $\tilde{f}(x) = f(x)$ when x is in $B(x_o,\varepsilon)$.

iii) $\lim\limits_{\|x\| \to \infty} f(x) = +\infty$.

Now, if x_o is a local minimum of f on S, x_o becomes a global minimum of f^* on S. In order to isolate x_o, we may substitute

$$\tilde{\tilde{f}}: x \longrightarrow \tilde{f}(x) + \|x-x_o\|^2 \quad \text{for} \quad \tilde{f}(x) \ .$$

In a neighborhood of x_o, $\tilde{\tilde{f}}$ is differentiable at x whenever f is differentiable at x. Thus, is no trouble in the calculation of the generalized jacobian matrix and if $F = (f,f_1,\ldots,f_m)^t$ and $\tilde{\tilde{F}} = (\tilde{\tilde{f}},f_1,\ldots,f_m)^t$ then

$$\tilde{J}_S(F;x_o) = \tilde{J}_S(\tilde{\tilde{F}},x_o) \qquad [4]$$

So, one can safely regard x_o as a unique and strong global minimum of f on S and suppose that $\lim\limits_{\|x\| \to \infty} f(x) = +\infty$.

REFERENCES

[1] E.J. McShane. *Extension of range of functions*. Bulletin of the American Mathematical Society, 40 (1934) p.837-843.

[2] R.H. Martin. *Nonlinear Operators and Differential Equations in Banach Spaces*. John Wiley & Sons (1976).

[3] R.T. Rockafellar. *Convex Analysis*. Princeton University Press (1970).

[4] T. Caputti. *The relative generalized jacobian matrix in the subdifferential calculus*.

[5] T. Caputti. *The plenary hull of the generalized jacobian matrix and the inverse function theorem in the subdifferential calculus*. Rev. UMA. Vol.30 - N°3/4 - 1982-1983. pp.190-196.

[6] R.T. Rockafellar. *Directionally Lipschitzian functions and subdifferential calculus*. Proceedings of the London Mathematical Society - Vol.39(3), 1979, pp.331-355.

Departamento de Matemática
Facultad de Ciencias Exactas
Universidad de Buenos Aires.

Recibido en diciembre de 1983.
Versión final agosto de 1984.