Revista de la
Unión Matemática Argentina
Volumen 36, 1990.

ON THE ε–SUBDIFFERENTIAL OF A CONVEX FUNCTION, II

TELMA CAPUTTI

1. INTRODUCTION

In this note we present a generalization of the algorithm for minimizing a locally Lipschitz continuous function f: $R^n \to R^1$ that is not necessarily differentiable or convex proposed by Caputti [1].

We present a descent algorithm designed to locate stationary points of functions of the form

$$F = h \circ f$$

where f: $R^n \to R^m$ is differentiable and h: $R^m \to R$ is convex. This problem and techniques to solve it play a central role in contemporary studies in mathematical programming. For example, the function h may be taken to be the identity, a norm, a penalty function or the distance function to some convex set. The conditions under which accumulation points of sequences generated by our algorithm are also stationary points of F can be determined via the notion of epi-convergence as the natural and appropriate technique.

2. THE ALGORITHM

Following Wets [2] we consider a mapping ρ: $R^n \to (-\infty,+\infty]$ that satisfies the following four conditions

i) ρ is a closed proper convex function

ii) $0 \in \text{int}(\text{dom } \rho)$

iii) $0 = \rho(0) = \min\{\rho(d) : d \in R^n\}$ (2.1)

iv) ρ is inf-compact, or equivalently, $\lim_{\|x\| \to \infty} \rho(x) = +\infty$.

We now employ these functions ρ in defining our primary analytic tool for the development of techniques intended to minimize (1.1) that is, the class of convex functions

$$d \longrightarrow \phi(d;x,\rho) : R^n \longrightarrow R$$

defined by the relation

$$\phi(d;x,\rho) := h(f(x)+f'(x)d) + \rho(d) \tag{2.2}$$

for every $x \in R^n$ and ρ as in the previous conditions (2.1) where $h: R^n \to R^1$ is a finite-valued convex function on R^m and $f: R^n \to R^m$ is Fréchet differentiable on R^n.

Taking into account [1, Proposition 2.2, Proposition 2.3] we may generalize as follows.

PROPOSITION 2.1. *Under the above conditions, the following statements are valid*

i) $0 \leqslant F(x) - \inf\{\phi(d;x,\rho) : d \in R^n\} \leqslant \varepsilon$ *if and only if* $0 \in \partial_\varepsilon \phi(0;x,\rho)$

ii) *If* $0 \notin \partial_\varepsilon \phi(0;x,\rho)$ *and* d *is any vector such that* $\psi^*(^d/\partial_\varepsilon \phi(0;x,\rho)) < 0$ *then* $F(x) - \inf_{\lambda > 0} \phi(\lambda d;x,\rho) > \varepsilon$ *where* ψ^* *denotes the support function of a non-empty compact convex subset* $\partial_\varepsilon \phi(0;x,\rho)$ [1].

Since this result is a straight-forward application of Proposition 2.2 and (2.3) in [1, Section 2] we omit its proof.

REMARK 2.2. The choice of search direction in the descent algorithm presented by Caputti [1, Section 3] can be generalized as follows.

For $x \in R^n$ and $r \in (0,1)$ we consider

$$\varepsilon(x) = \begin{cases} 0 & \text{if} \quad 0 \in \partial_{\varepsilon}\phi(0;x,\rho) \\ \max\{r^p: 0 \notin \partial_{r^p}\phi(0;x,\rho),\ p = 0,1,\ldots\} & \text{otherwise.} \end{cases} \tag{2.3}$$

and we define the class of search directions by

$$D(x;\rho,r) = \begin{cases} \{0\} & \text{if} \quad \varepsilon(x) = 0 \\ \{d: \psi^*(d/\partial_{\varepsilon(x)}\phi(0;x,\rho) < 0\} \text{ and } F(x) - \phi(d;x,\rho) \geqslant \varepsilon(x) & \text{otherwise.} \end{cases} \tag{2.4}$$

If $0 \notin \partial\phi(0;x,\rho)$ the statement that the set $D(x;\rho,r)$ is non-empty is easily seen to be equivalent to the above Proposition 2.1, part ii).

ALGORITHM 2.2. The type of algorithm that we study is of the form

$$x_{k+1} = x_k + \lambda_k d_k \tag{2.5}$$

where

$$\lambda_k = \max\{r^p: F(x_k) - F(x_k + r^p d_k) \geqslant r^p \varepsilon(x),\ p = 0,1,2,\ldots\},$$

$d_k \in D(x;\rho,r)$, $\varepsilon(x) \geqslant 0$, $r \in (0,1)$ and $x \in R^n$.

We note that the number $\varepsilon(x)$ and the set $D(x;\rho,r)$ satisfy the following three conditions:

i) $D(x;\rho,r) \neq \emptyset$ according with the above observation

ii) $0 \in D(x;\rho,r)$ if and only if $\varepsilon(x) = 0$ if and only if $0 \in \partial F(x_k)$ (2.6)

iii) $h(f(x_k) + f'(x_k)d_k) - F(x_k) \leqslant -\varepsilon(x)$

where $\partial F(x_k)$, the Clarke subdifferential of F at x_k, has the representation

$$\partial F(x_k) = \partial h(f(x_k)) \circ f'(x_k) = \{y \in R^n / y = zf'(x_k),\ z \in \partial h(f(x_k))\}$$

for all $x \in R^n$ [3, Proposition 10].

We see that any direction d_k for which $h(f(x_k) + f'(x_k)d_k) < h(f(x_k))$ is a descent direction for F. Therefore, conditions (2.6) along with the stopping criterion $0 \in \partial F(x_k)$ (x_k

is a stationary point of F) guarantee that algorithm (2.5) is always well defined.

The convergence result for algorithms of type (2.5) that satisfy conditions (2.6) to the effect that,

"If $\{x_k\}$ is the sequence generated by algorithm (2.5) with initial point x_0 and stopping criterion $0 \in \partial F(x)$, then provided that conditions (2.6) hold, one of the following must occur: the algorithm terminates finitely at x_{k_0} with $0 \in \partial F(x_{k_0})$; and/or $F(x_k) \downarrow -\infty$; and/or the sequence $\{\|d_k\|\}$ diverges to $+\infty$", is similar to the proof of convergence given by Bertsekas and Mitter in their paper [4].

We may view the algorithm as successively minimizing a sequence of convex functions that are themselves local approximations to the function in which our real interest lies. The natural and appropriate technique by which such optimization schemes are analyzed is via the notion of epi-convergence. The basic properties of epi-convergent sequences of convex functions as applied to optimization problems, are developed in, for example,the work of Rockafellar and Wets [2,5].

DEFINITION 2.3. Let $\{f_i\}_{i=0}^{\infty}$ be a sequence of closed convex functions with domain in R^n and range $R^* := R \cup \{+\infty\}$. We say that $\{f_i\}$ converges pointwise to the closed convex function $f: R^n \to R^*$ and write $f_i \overset{p}{\to} f$ if $\lim_i f_i(x) = f(x)$ for all $x \in R^n$. We say that $\{f_i\}$ epi-converges to f and write $f_i \overset{e}{\to} f$ if the epi-graphs of the f_i converge to the epi-graph of f, that is

$$\lim_i \sup \operatorname{epi}(f_i) = \lim_i \inf \operatorname{epi}(f_i)$$

where the epi-graph of a convex function $g: R^n \to R^*$ is the set

$$\operatorname{epi}(g) = \{(x,\alpha) \in R^n \times R: g(x) \leqslant \alpha;\ x \in \operatorname{dom}(g)\} .$$

Taking into account [2, Corollary 4, Theorem 7, Theorem 9] we may prove the following result

THEOREM 2.4. *Let* $\{x_k\}$ *be the sequence generated by Algorithm 2.5 with initial point* $x_0 \in R^n$*, stopping criterion* $0 \in \partial F(x)$. *If* x^* *is an accumulation point of* $\{x_k\}$ *with* $y_j \underset{j}{\to} x^*$ *and* $\rho(.,y_j) \overset{p}{\underset{j}{\to}} \rho(.,x^*)$ *for some subsequence* $\{y_j\}$ *of* $\{x_k\}$*, then*

$$\underset{j}{\mathrm{Lim}}(\min\{\phi(d;y_j,\rho):\ d \in R^n\}) = \min\{\phi(d;x^*,\rho):\ d \in R^n\} = F(x^*)\ , \qquad (2.7)$$

$F(x_k) \downarrow F(x^*)$ *and* $0 \in \partial F(x^*)$.

Proof. Let $\{y_j\}$ be as in the hypothesis with $\rho(.,y_j) \overset{p}{\to} \rho(.,x^*)$. Then, $\phi(.;y_j,\rho) \overset{p}{\to} \phi(.;x^*,\rho)$. Hence, by [2, Corollary 4], $\phi(.;y_j,\rho) \overset{e}{\to} \phi(.;x^*,\rho)$. Thus, by [2, Theorem 7], we have the first half of (2.7) and by [2, Theorem 9], (2.7) also holds. Furthermore, $F(x_k) \downarrow F(x^*)$, since $\{F(x_k)\}$ is a decreasing sequence and the sequence $\{\|d_k\|\}$ is uniformly bounded since $\{d_i\} \subset \{d:\phi(d;x,\rho) \leqslant F(x)\}$.

From Proposition 2.1 i) we know that

$$0 \leqslant F(x_k) - \min\{\phi(d;x,\rho):\ d \in R^n\} \leqslant r^{-1}\ \varepsilon(x)$$

for all k sufficiently large.

We have that,

$$\underset{k}{\mathrm{Lim}}\ [F(x_k) - \min\{\phi(d;x_k,\rho):\ d \in R^n\}] = 0$$

Therefore, by the first half of (2.7), we have that

$$F(x^*) = \underset{j}{\mathrm{Lim}}\ F(y_j) = \underset{j}{\mathrm{Lim}}\ [\ \min\{\phi(d;y_j,\rho):\ d \in R^n]\ =$$
$$= \min\{\phi(d;x^*,\rho):\ d \in R^n\}$$

and so $0 \in \partial F(x^*)$.

(q.e.d.)

REFERENCES

[1] CAPUTTI,T., *On the ε-subdifferential of a convex function*, Rev. Unión Mat. Argentina. Volumen 31, 1984, p.p.202-210.

[2] WETS,R.J.-B., *Convergence of convex functions, variational inequalities and convex optimization problems*, in R.W. Cottle, F.Giannessi and J.H.Lions, eds., Variational inequalities and complementary problems. John Wiley and Sons, New York, 1980, p.p. 275-403.

[3] CLARKE,F.H., *A new approach to Lagrange multipliers*, Mathematics of Operations Research, Volume 1, Nro.2, May, 1976, p.p.165-174.

[4] D.P.BERTSEKAS-S.K.MITTER, *A descent numerical method for optimization problems with nondifferentiable cost functionals*, SIAM Journal of Control and Optimization 11, 1973. p.p.637-652.

[5] ROCKAFELLAR,R.T., *Convex Analysis*, Princeton University Press Princeton. New Jersey, 1970.

Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires.

Recibido en junio de 1990.