Report n.26

Invexity in Nonsmooth Programming

Giorgio GIORGI* - Stefan MITITELU**

Pisa, 1989

* Dipartimento di Ricerche Aaziendali - Università di Pavia
** Institute of Civil Engineering - Bucharest

INVEXITY IN NONSMOOTH PROGRAMMING

Giorgio Giorgi (o)
Stefan Mititelu (oo)

Summary. The aim of this paper is to show that the class of the invex lipschitzian functions is equivalent to the class of functions whose stationary points are global minima. Moreover, we shall generalize the Fritz John and the Kuhn-Tucker necessary optimality conditions for a nonlinear programming problem (P) whose lagrangian is an invex lipschitzian function. Under the same hypothesis on (P), we shall generalize the direct duality theorem of Wolfe and the strict converse duality theorem of Mangasarian.

§1. *Introduction*.

In 1975 Clarke (2) introduced the notion of *generalized directional derivative* (Clarke derivative) of a locally lipschitzian function (see Definition 2.1) $f: X \to R$ (X is a Banach space) at a point $x \in X$, in the direction $h \in X$, in the following way:

$$f^{o}(x;\, h) = \limsup_{\substack{x' \to x \\ \lambda \downarrow 0}} \frac{f(x' + \lambda h) - f(x')}{\lambda} .$$

It can be shown that $f^{o}(x;\, .)$ is, for each $x \in X$, a finite, sublinear (i. e. convex and positively homogeneous) function of h. From these properties Clarke defines the *generalized gradient* of $f(x)$ (Clarke subdifferential of $f(x)$), denoted

(o) Dipartimento di Ricerche Aziendali - Università di Pavia.
(oo) Institute of Civil Engineering - Bucharest.

$\partial f(x)$, as the subdifferential of the convex function $f^o(x; .)$ at O, that is

$$\partial f(x) = \{\xi \in X^* | f^o(x; h) \geq < \xi, h >, \forall h \in X\} ,$$

where X^* is the dual space of X.

The set $\partial f(x)$ is a nonempty, convex and w^*-compact subset of X^* and satisfies the relation

$$f^o(x; h) = \max \{ < \xi, h > | \xi \in \partial f(x) \}, \text{ for all } h \in X.$$

Clarke subdifferential keeps many properties of the differential, as shown in the papers of Clarke (2, 3), Hiriart-Urruty (9, 10), Rockafellar (16). Clarke subdifferential found an ideal place of application in the optimization theory; in particular this concept was applied in the study of the necessary Fritz John and Kuhn-Tucker necessary optimality conditions (3, 9, 10, 16) and of the Wolfe duality theorem (5) for a nonsmooth programming problem involving locally lipschitzian functions.

In the present paper we prove first that the class of the invex locally lipschitzian functions is equivalent to the class of functions whose stationary points (in the sense of Clarke) are global minima. This property is then used to establish some optimality theorems and some duality theorems, in the Wolfe approach, for a nonsmooth programming problem.

§2. *Invexity for lipschitzian functions.*

The notion of invexity was introduced (without this name) by Hanson (8) for Fréchet differentiable functions and was developed mostly by Hanson, Mond and Craven. The name of "invex functions" was given by Craven (4) as a contraction of "invariant convex", since $f = g \circ \theta$ will be invex if g is convex, θ is differentiable and $\nabla\theta$ has full rank: invexity (unlike convexity) is a property invariant to bijective coordinate transformations. Let $X \subseteq R^n$ an open set and $f: X \to R$ be differentiable; then f(x) is *invex* if

(1) $$f(x) - f(u) \geqq [h(x, u)]^T \nabla f(x),$$

for all $x, u \in X$ and for some vector function $h: X \times X \to R^n$.

Of course if $h(x, u) = x - u$, we obtain from (1) the classical definition of differentiable convex function. Craven and Glover (6) showed that the class of the invex functions is equivalent to the class of functions whose stationary points are global minima (see also (1, 11)). We shall prove that this property continues to be true for invex locally lipschitzian functions. Let us give the following definitions.

Definition 2.1.

Let $X \subseteq R^n$; $f: X \to R$ is said to be *locally lipschitzian* if for any point $x \in X$ there exists a neighbourhood N of x such that, for some nonnegative constant k and for all $z, u \in N$, the following relation holds

$$|f(z) - f(u)| \leqq k \|z - u\|.$$

As previously noted, under this property $f^o(x; h)$ is a finite and sublinear function of h.

Definition 2.2.

Let $X \subseteq R^n$ and $f: X \to R$ be locally lipschitzian; f(x) is *L-invex at* $u \in X$ if there exists a vector function $\eta: X \times X \to R^n$ such that

(2) $f(x) - f(u) \geqq [\eta(x, u)]^T \xi(u), \forall \xi(u) \in \partial f(u), \forall x \in X$

or equivalently (5)

$$f(x) - f(u) \geqq f^{o}(u, \eta(x, u)), \forall x \in X.$$

The function f(x) is *L-strictly invex at u ∈ X* if, for $x \neq u$, the relation (2) is satisfied with the strict inequality sign. The function f(x) is *L-invex on X (L-strictly invex on X)* if it is L-invex (L-strictly invex) at any point of X.

Definition 2.3.

Any point $x \in X \subseteq R^n$ which satisfies the condition

$$0 \in \partial f(x)$$

is called *Clarke stationary point of f(x)* (7).

The function $\eta(x, u)$ can be chosen in the following way:

(3) $$\eta(x, u) = \begin{cases} 0, \text{ if } 0 \in \partial f(u) \\ \min\limits_{\xi(u) \in \partial f(u)} \dfrac{f(x) - f(u)}{[\xi(u)]^T \xi(u)} \xi(u), \\ \text{if } 0 \notin \partial f(u). \end{cases}$$

Example. The function $f(x) = |x|$, $x \in R$, is (locally) lipschitzian on R but not differentiable at $x = 0$. We have

$$\partial f(u) = \begin{cases} 1 \quad, \text{ if } u > 0 \\ -1 \quad, \text{ if } u < 0 \end{cases}$$

and

$$\partial f(0) = \{\xi_o \mid \xi_o = \lambda(-1) + (1 - \lambda)\, 1,\ 0 \leqq \lambda \leqq 1\} =$$

$$= \{\xi_o \mid \xi_o = 1 - 2\lambda,\ 0 \leqq \lambda \leqq 1\}.$$

So we deduce

$$\eta(x, u) = \begin{cases} -|x| - u, & \text{if } u < 0 \\ -|x|, & \text{if } u = 0 \\ |x| - u, & \text{if } u > 0. \end{cases}$$

The following theorem generalizes the above mentioned result of Craven and Glover.

Theorem 2.1. The lipschitzian function f is L-invex on X if and only if every Clarke stationary point is a point of global minimum on X.

Proof. If f(x) is L-invex, then $0 \in \partial f(x)$ implies $f(x) - f(u) \geqq 0$, $\forall\, x \in X$, that is u is a global minimum point.

Conversely, let us suppose that

$$0 \in \partial f(u) \Longrightarrow f(x) \geqq f(u), \ \forall\, x \in X.$$

Then the function $\eta(x, u)$ can be selected according to relation (3).

§3. *Necessary optimality conditions for a nonsmooth programming problem.*

Let us consider the locally lipschitzian functions

$$f(x),\ g_1(x),\ g_2(x),\ \ldots,\ g_m(x),$$

all from $X \subseteq R^n$ to R, where X is a nonempty open set, and the programming problem

$$\text{(P)} \qquad \begin{cases} \min\limits_{x \in D} f(x) \\ D = \{x \mid x \in X,\ g_i(x) \leqq 0,\ i = 1,2,\ldots,m\}. \end{cases}$$

Let us suppose that $x^o \in D$ is a point of local minimum of (P); Clarke (3) establishes the following Fritz John necessary optimality conditions for (P): there exists numbers $r^o, u_1^o, \ldots, u_m^o \in R$, not all zero, such that

(4) $$r^o \, \partial f(x^o) + \sum_{i=1}^{m} u_i^o \, \partial g_i(x^o) \ni 0$$

(5) $$u_i^o \, g_i(x^o) = 0, \; i = 1,\ldots,m$$

(6) $$(r^o, u_i^o) \geqq 0, \; i = 1, \ldots, m.$$

Under various constraint qualifications, Clarke (3), Hiriart-Urruty (9), Watkins (18) established the following Kuhn-Tucker necessary optimality conditions for the point x^o, local solution of (P): there exists numbers $u_1^o,\ldots,u_m^o \in R$ so that

(7) $$\partial f(x^o) + \sum_{i=1}^{m} u_i^o \, \partial g_i(x^o) \ni 0$$

(8) $$u_i^o \, g_i(x^o) = 0, \; i = 1,\ldots, m$$

(9) $$u_i^o \geqq 0, \; i = 1,\ldots,m.$$

Let us introduce the notation

$$\Phi_r(x, u) = r \, f(x) + \sum_{i=1}^{m} u_i \, g_i(x) = r \, f(x) + u^T g(x).$$

The following theorem is a Fritz John saddle point necessary optimality theorem for problem (P).

Theorem 3.1. Let $x^o \in D$ be a solution of the program (P) and let us suppose that the functions $f, g_1, \ldots, g_m$ are locally lipschitzian on X. Let $r^o \in R$ and $u^o \in R^m$, so that the program (P) satisfies at (r^o, u^o, x^o) the Fritz John conditions (4)-(6). If $\Phi_{r^o}(x,u^o)$ is L-invex at x^o, then the following relations hold:

(10) $$\Phi_{r^o}(x^o, u) \leqq \Phi_{r^o}(x^o, u^o) \leqq \Phi_{r^o}(x, u^o), \ \forall x \in X, \ \forall u \geqq 0,$$

that is Φ_{r^o} admits at (x^o, u^o) a saddle point on $X \times R^m_+$.

Proof. Let first note that the relation (4) can be rewritten with the following notation

(11) $$\partial_x \Phi_{r^o}(x^o, u^o) \ni 0.$$

But the function $\Phi_{r^o}(x, u^o)$ being L-invex at x^o, according to theorem 2.1 from (11) it results

(12) $$\Phi_{r^o}(x^o, u^o) \leqq \Phi_{r^o}(x, u^o), \ \forall x \in X.$$

The inequality

(13) $$\Phi_{r^o}(x^o, u) \leqq \Phi_{r^o}(x^o, u^o), \ u \in R^m_+$$

is equivalent to the relation

$$u^T g(x^o) \leqq 0, \ \forall u \geqq 0.$$

Thus the relations (12) and (13) are equivalent to (10).

Moreover it can be proved the following Kuhn-Tucker saddle point necessary optimality conditions for the problem (P).

Theorem 3.2. Let $x^o \in D$ be a solution of program (P) and let us suppose that the functions $f, g_1,\dots,g_m$ are locally lipschitzian on X. We also suppose that the program (P) satisfies at x^o a constraint qualification so that the vector

(x^o, u^o) verifies the Kuhn-Tucker conditions (7)-(9). If the function $\Phi(x, u^o)$ is L-invex at x^o, then the following inequalities hold:

$$\Phi(x^o, u) \leqq \Phi(x^o, u^o) \leqq \Phi(x, u^o),\ \forall\, x \in X,\ \forall\, u \in R^m_+.$$

The proof is similar to the one of the previous theorem (use the Kuhn-Tucker conditions (7)-(9) instead of the Fritz John conditions (4)-(6)).

§ 4. *Wolfe duality.*

The Wolfe dual of the program (P), according to Schechter (17), is the following program

(DW)
$$\begin{cases} \max\limits_{(x,u)\in\Omega} \Phi(x, u) \\ \\ \Omega = \{(x, u) \in R^{n+m} \mid \partial_x \Phi(x, u) \ni 0,\ x \in X,\ u \geqq 0\}. \end{cases}$$

Theorem 4.1.(Weak Wolfe duality theorem). Let us suppose D and Ω nonempty and that the function $\Phi(., u)$ is L-invex on X for each $u \geqq 0$. Then

(14)
$$\inf_{x \in D} f(x) \geqq \sup_{(z,u)\in\Omega} \Phi(z, u).$$

Proof. For $(z, u) \in \Omega$ we have

(15)
$$\partial_x \Phi(z, u) \ni 0.$$

Being $\Phi(., u)$ L-invex on X, according to theorem 2.1, it results

$$\Phi(x, u) \geqq \Phi(z, u),\ \forall\, x \in X,$$

or equivalently

(16)
$$f(x) + u^T g(x) \geqq \Phi(z, u),\ \forall\, x \in X.$$

But $x \in D$ and $u \geqq 0$ involves $u^T g(x) \leqq 0$. Then from (16) it results

(17) $$f(x) \geqq \Phi(z, u), \ \forall x \in D;$$

(z, u) being arbitrary in Ω, from (17) we obtain the thesis.

Theorem 4.2. (Direct Wolfe duality theorem). Let x^o be a solution of the primal program (P) and let $f, g_1, ., g_m$ be locally lipschitzian functions on X; moreover let the following conditions hold:

(1) A constraint qualification for (P) is satisfied.

(2) The function $\Phi(., u)$ is L-invex on X, for each $u \geqq 0$.

Then there exists $u^o \in R^m$ so that the point (x^o, u^o) is a solution of the dual program (DW) and it is $\Phi(x^o, u^o) = f(x^o)$.

Proof. Since x^o is a solution of (P) and a constraint qualification is satisfied, there exists a point $u^o \in R^m$ satisfying the Kuhn-Tucker conditions (7)-(9). So, $(x^o, u^o) \in \Omega$ and $(u^o)^T g(x^o) = 0$. As the function $\Phi(., u)$ is L-invex on X, the relation (17) comes true. Then for $x = z = x^o$ and $u = u^o$, from (17) we deduce the equality

$$\min_{x \in D} f(x) = \max_{(z,u) \in \Omega} \Phi(z, u) = \Phi(x^o, u^o) = f(x^o).$$

Theorem 4.3 (Strict converse Wolfe duality theorem). Let

Let (x^o, u^o) be a solution of the dual program (DW) and let the following conditions be satisfied:

(1) The primal program (P) admits a solution $\bar{x}$ and a constraint qualification is satisfied for the same program.

(2) The function $\Phi(x, u^o)$ is L-strictly invex at x^o. Then $x^o = \bar{x}$ and $f(x^o) = \Phi(x^o, u^o)$.

The proof is similar to the one of theorem 4 of (14) in which the Kuhn-Tucker conditions (7)-(9) are to be used.

The theorem 4.3 admits a local version. Indeed, let $V \subset X$ a neighbourhood of the point x^o and consider the set $B = (V \cap R^m_+) \times \Omega$. Then the following corollary holds.

Corollary. Let (x^o, u^o) be a solution on B of the program (DW); suppose moreover that the following conditions hold:

(1) The primal program (P) admits the solution $\bar{x}$ on $V \cap D$, and a constraint qualification is satisfied for (P).

(2) The function $\Phi(x, u^o)$ is strictly L-invex at x^o. Then $x^o = \bar{x}$ (that is x^o is a local solution of the primal program (P)) and $f(x^o) = \Phi(x^o, u^o)$.

Remark. The theorem 4.2 generalizes the direct duality theorem of Wolfe (19) and the theorem 4.3 generalizes the strict converse duality theorem of Mangasarian (12).

REFERENCES

(1) A. Ben-Israel, B. Mond, *What is invexity?*, J. Austral. Math. Soc. Ser. B, 28, (1986), 1-9.

(2) F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc., 205 (1975), 247-262.

(3) F. H. Clarke, *A new approach to Lagrange multipliers*, Math. Oper. Res., 1, 2 (1976), 165-174.

(4) B. D. Craven, *Duality for generalized convex fractional programs*, in: S. Schaible, W. T. Ziemba (Eds.),

Generalized concavity in optimization and economics, Academic Press, New York, 1981, 473-489.

(5) B. D. Craven, *Nondifferentiable optimization by smooth approximation*, Optimization, 17, (1986), 3-17.

(6) B. D. Craven, B. M. Glover, *Invex functions and duality*, J. Austral. Math. Soc. Ser. A, 39, (1985), 1-20.

(7) V. F. Demyanov, L. N. Polyakova, A. M. Rubinov, *Nonsmoothness and quasidifferentiability*, Math. Programming Study 29, 1986, 1-19.

(8) M. A. Hanson, *On sufficiency of the Kuhn-Tucker conditions*, J. Math. Anal. Appl., 80, (1981), 545-550.

(9) J. B. Hiriart-Urruty, *On optimality conditions in nondifferentiable programming*, Math.Programming, 14, (1978), 73-86.

(10) J. B. Hiriart-Urruty, *Refinements of necessary optimality conditions in nondifferentiable programming I*, Appl. Math. Optim., 5, (1979), 63-82.

(11) V. Jeyakumar, *Strong and weak invexity in mathematical programming*, Methods Oper. Res., 55, (1985), 109-125.

(12) O. L. Mangasarian, *Duality in nonlinear programming*, Quart. Appl. Math., 20, (1962), 300-303.

(13) O. L. Mangasarian, *Nonlinear programming*, Mc Graw-Hill, New York, 1969.

(14) S. Mititelu, *Dualitatea Wolfe fără convexitate*, Studii și cercetari matematice, 38, (1986), 302-307.

(15) S. Mititelu, *Optimality, minimax and duality in nonlinear*

programming, Optimization, 18, (1987), 501-506.

(16) R. T. Rockafellar, *La théorie des sous-gradients et ses applications à l'optimisation*, Les Presses de l'Université de Montréal, Montréal, 1979.

(17) M. Schechter, *A subgradient duality theorem*, J. Math. Anal. Appl., 61, (1977), 850-855.

(18) G. G. Watkins, *Nonsmooth Milyutin-Dubovitskii theory and Clarke tangent cone*, Math. Oper. Res.,12, (1986), 70-80.

(19) P. Wolfe, *A duality theorem for nonlinear programming*, Quart. Appl. Math., 19, (1961), 239-244.