Control and Optimization in Applied Mathematics (COAM)
Vol. 7, No. 1, Winter-Spring 2022 (79-92), ©2016 Payame Noor University, Iran

DOI. 10.30473/COAM.2022.63088.1196 (Cited this article)

Research Article

Linearization and Gap Function in Nonsmooth Quasiconvex Optimization Using Incident Subdifferential

Hamed Soroush*

Department of Mathematics, Payame Noor University (PNU),
P.O. Box. 19395-4697, Tehran, Iran.

Received: February 18, 2022; **Accepted:** June 14, 2022.

Abstract. The purpose of this paper is to develop nonsmooth optimization problems (P) in which all emerging functions are assumed to be real-valued quasiconvex functions that are defined on a finite-dimensional Euclidean space. First, we introduce two linear optimization problems with the same optimal value of the considered problem. Then, we introduce a real-valued non-negative gap function for (P), and we provide some conditions which ensure that its null points are the same as the optimal solution of problem (P). The results are based on incident subdifferential, which is an important concept in the analysis of quasiconvex functions.

Keywords. Quasiconvex optimization, Linearization, Gap function, Incident subdifferential.

MSC. 90C34; 90C40.

* Corresponding author
h_soroush2011@pnu.ac.ir
http://mathco.journals.pnu.ac.ir

1 Introduction

In this paper, we consider the following *quasiconvex optimization problem*:

$$\begin{aligned} (\text{P}): \qquad & \inf \quad f(x) \\ & \text{s.t.} \quad g_j(x) \leq 0, \quad j \in J, \\ & \qquad\ \ x \in \mathbb{R}^n, \end{aligned}$$

where, f and g_j, for $j \in J$, are real-valued quasiconvex functions that are defined on $\mathbb{R}^n$, and J is a finite index set. Problems of this kind have been utilized for the analysis of topological optimization problems and other theoretical aspects (see e.g., [10, 11, 14, 15, 16] and the references therein). Gap function for a mathematical programming problem has been studied in various publications. Hearn [8] introduced the gap function for a differentiable convex optimization problem in finite-dimensional Euclidean spaces, for the first time. For a thorough study of this subject, refer to the works of Auslender [3], Altangerel et al. [1, 2], Chen et al. [5], Hasani and Sadeghieh [7], Kanzi et al. [10], and the recent paper by Caristi et al. [4]. One of the basic properties of gap function for optimization problems is its ability in characterizing the solutions of the problem in question. On the other hand, the linearization of nonlinear, especially nonsmooth, optimization problems is an important topic in optimization theory; see e.g., [11, 13]. Linearization methods can be used to convert a nonlinear optimization problem into a linear optimization problem. In this process, extra variables and constraints are introduced to construct the original problem. Various methods have been proposed in the literature by linearizing a nonlinear problem [12, 19]. The aim of the present paper is to introduce and examine the linearization method as well as a suitable gap function for problem (P). Since we do not assume that the functions which appear in the problem are differentiable, we should state our results by a suitable subdifferential. As it is shown in [14] and [16, Section 5], the incident subdifferential is an important subdifferential in the analysis of quasiconvex functions. Very recently, the necessary and sufficient optimality conditions and also the weak and strong duality results for problem (P) have been presented according to incident subdifferential in [18]. Since the linearization and also the characterization of solutions of (P) with gap function have not yet been studied by incident subdifferential, we will present our results according to this subdifferential to fill this gap.

The structure of subsequent sections of this paper is as follows. In Section 2, we present the required notations, definitions and preliminary results from quasiconvex analysis and optimization theory. Section 3, contains the main results of the paper, that are based on the results of the very recent paper [18]. Conclusion is expressed in Section 4.

2 Notations and Preliminaries

In this section we present some definitions, notations, and auxiliary results that will be needed in the sequel. For a given $S \subseteq \mathbb{R}^n$, the superfluous convex cone hull of S is denoted by $cone(S)$. Also, if $S \neq \emptyset$ is a subset of $\mathbb{R}^n$ and $\hat{x} \in S$, the attainable cone, the contingent cone, and the interior cone of S at $\hat{x} \in S$ are respectively defined by ([6])

$$\mathcal{A}(S,\hat{x}) := \Big\{x \in \mathbb{R}^n \mid \forall t_\ell \downarrow 0, \exists x_\ell \to x, \hat{x} + t_\ell x_\ell \in S,\ \forall \ell \in \mathbb{N}\Big\},$$

$$\mathcal{Z}(S,\hat{x}) := \Big\{x \in \mathbb{R}^n \mid \exists t_\ell \downarrow 0, \exists x_\ell \to x, \hat{x} + t_\ell x_\ell \in S,\ \forall \ell \in \mathbb{N}\Big\},$$

$$\mathcal{I}(S,\hat{x}) := \Big\{x \in \mathbb{R}^n \mid \exists L > 0,\ \forall t_\ell \downarrow 0, \forall x_\ell \to x, \hat{x} + t_\ell x_\ell \in S,\ \forall \ell \geq L\Big\}.$$

Obviously, the following inclusions hold

$$\mathcal{I}(S,\hat{x}) \subseteq \mathcal{A}(S,\hat{x}) \subseteq \mathcal{Z}(S,\hat{x}) \subseteq conv\Big(\mathcal{Z}(S,\hat{x})\Big).$$

A substantial list of equivalent forms, examples, properties, and details of these cones can be found in [6, Section 3.4]. We recall from [16] that the incident (or upper epi-) directional derivative of a given function $\varphi : \mathbb{R}^n \to \mathbb{R}$ at $\hat{x} \in \mathbb{R}^n$ in direction $d \in \mathbb{R}^n$, and the incident subdifferential of φ at $\hat{x}$ are, respectively, defined by

$$\varphi^\flat(\hat{x};d) := \sup_{\delta>0} \limsup_{\varepsilon\downarrow 0} \inf_{\|w-d\|<\delta} \frac{\varphi(\hat{x}+\varepsilon w) - \varphi(\hat{x})}{\varepsilon},$$

and

$$\partial^\flat \varphi(\hat{x}) := \Big\{\xi \in \mathbb{R}^n \mid \langle \xi, v\rangle \leq \varphi^\flat(\hat{x};v), \quad \forall v \in \mathbb{R}^n\Big\},$$

where, $\langle \cdot,\cdot \rangle$ denotes the standard inner product in $\mathbb{R}^n$. Also, we say that:

- φ is regular at $\hat{x}$ if the function $d \to \varphi^\flat(\hat{x};d)$ is convex.
- φ is quasiconvex if for all x_1 and x_2 in $\mathbb{R}^n$ and for all λ in $[0,1]$ the following inequality holds.

$$\varphi\Big(\lambda x_1 + (1-\lambda)x_2\Big) \leq \max\Big\{\varphi(x_1), \varphi(x_2)\Big\}.$$

Some extraordinary properties for $\varphi^\flat(\hat{x};d)$ and $\partial^\flat \varphi(\hat{x})$ that show their important role in quasiconvex analysis are presented in [14, Propositions 1, 3, 16]. We can see ([14, Proposition 47, and Corollary 50]) that if $\psi : \mathbb{R}^n \to \mathbb{R}$ is a convex function and $\hat{x} \in \mathbb{R}^n$, then

$$\partial^\flat \psi(\hat{x}) = \partial\psi(\hat{x}) := \Big\{\xi \in \mathbb{R}^n \mid \psi(x) - \psi(\hat{x}) \geq \Big\langle \xi, x - \hat{x}\Big\rangle, \quad \forall\, x \in \mathbb{R}^n\Big\}.$$

The following theorem has a key role in convex optimization.

Theorem 1. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a convex function. If the minimum of ϕ on a convex set $B \subseteq \mathbb{R}^n$ is attained at $\hat{x} \in B$, one has

$$0_n \in \partial\phi(\hat{x}) + N(B,\hat{x}),$$

where, 0_n denotes the zero vector in $\mathbb{R}^n$, and $N(B,\hat{x})$ denotes the normal cone of B at $\hat{x}$, defined by

$$N(B,\hat{x}) := \Big\{x \in \mathbb{R}^n \mid \langle x, b - \hat{x}\rangle \leq 0, \quad \forall b \in B\Big\}.$$

It is easy to see that if the function $\phi : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable at $\hat{x} \in \mathbb{R}^n$, then $\partial^b \phi(\hat{x}) = \{\nabla\phi(\hat{x})\}$, where $\nabla\varphi(\hat{x})$ denotes the gradient of ϕ at $\hat{x}$, and $\phi'(\hat{x};d) = \langle d, \nabla\phi(\hat{x})\rangle$, where $\phi'(\hat{x};d)$ denotes the classic directional derivative of ϕ at $\hat{x}$ in direction $d \in \mathbb{R}^n$, defined by

$$\varphi'(\hat{x};d) := \lim_{\varepsilon\to 0} \frac{\varphi(\hat{x}+\varepsilon d) - \varphi(\hat{x})}{\varepsilon}.$$

As the final point of this section, we recall from the classical theory of optimization that if ϑ and θ_i, as $i = 1,\ldots,m,$ are continuously differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}$, we say that the optimization problem

$$\begin{array}{lll} (\Delta): & \inf & \vartheta(x) \\ & \text{s.t.} & \theta_i(x) \le 0, \quad i = 1,\ldots,m, \\ & & x \in \mathbb{R}^n, \end{array}$$

satisfies the Guignard constraint qualification (GCQ) at $x^0 \in Q$, with

$$Q := \{x \in \mathbb{R}^n \mid \theta_i(x) \le 0, \quad i = 1,\ldots,m\},$$

when

$$\Big\{d \in \mathbb{R}^n \mid \langle\nabla\theta_i(x^0), d\rangle \le 0, \quad i \in I^*\Big\} \subseteq conv\Big(\mathcal{Z}(Q, x^0)\Big),$$

where, $I^* := \Big\{i = 1,\ldots,m \mid \theta_i(x^0) = 0\Big\}$. The following theorem, states the classical Karush-Kuhn-Tucker (KKT) necessary optimality condition for problem (Δ).

Theorem 2. [6, Theorem 3.6.3] Assume that x^0 is an optimal solution of problem (Δ) and GCQ holds at x^0. Then, there exist non-negative scalars λ_i, as $i \in I^*$, such that

$$\nabla\vartheta(x^0) + \sum_{i\in I^*} \lambda_i \nabla\theta_i(x^0) = 0_n.$$

Note that GCQ is the weakest constraint qualification that concludes the above equality that is named KKT relation; see, e.g., [6, Page 279].

3 The Main Results

As a starting point of this section, we mention that the feasible set of (P) is denoted by Ω, i.e.,

$$\Omega := \{x \in \mathbb{R}^n \mid g_j(x) \le 0 \qquad \forall j \in J\}.$$

Also, for each $\hat{x} \in \Omega$, the set of active indices at $\hat{x}$ is defined by

$$J(\hat{x}) := \Big\{j \in J \mid g_j(\hat{x}) = 0\Big\}.$$

The following constraints qualification and KKT type necessary optimality conditions can be found in [18].

Definition 1. [18, Definition 1] We say that the generalized Kuhn-Tucker constraints qualification (GKTCQ, in brief) holds at $\hat{x} \in \Omega$ when

$$\Big\{d \in \mathbb{R}^n \mid \langle d, \xi \rangle \leq 0, \quad \forall \xi \in \bigcup_{j \in J(\hat{x})} \partial^\flat \psi_j(\hat{x})\Big\} \subseteq \mathcal{A}(\Omega, \hat{x}).$$

Theorem 3. (KKT necessary conditions)[18, Theorem 5] Let $\hat{x} \in \Omega$ be an optimal solution of (P), ϑ and ψ_j be regular at $\hat{x}$ for $j \in J(\hat{x})$, and $\mathcal{I}(\Omega, \hat{x}) \neq \emptyset$. Moreover, assume that GKTCQ is satisfied at $\hat{x}$ and the following cone is closed:

$$cone\Big(\bigcup_{j \in J(\hat{x})} \partial^\flat \psi_j(\hat{x})\Big).$$

Then, there exist $\lambda_j \geq 0$ for $j \in J(\hat{x})$ such that

$$0_n \in \partial^\flat \vartheta(\hat{x}) + \sum_{j \in J(\hat{x})} \lambda_j \partial^\flat \psi_j(\hat{x}), \qquad and \qquad \sum_{j \in J(\hat{x})} \lambda_j = 1.$$

The following important theorem will be used in the sequel.

Theorem 4. [18, Lemma 2] Suppose that $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a quasiconvex function. Then,

$$\varphi(x) \leq \varphi(\hat{x}) \Longrightarrow \langle \xi, x - \hat{x} \rangle \leq 0, \qquad \forall \xi \in \partial^\flat \varphi(\hat{x}).$$

Equivalently, the quasiconvexity of φ and the inequality $\langle \xi, x - \hat{x} \rangle > 0$, for some $\xi \in \partial^\flat \varphi(\hat{x})$, conclude that $\varphi(x) > \varphi(\hat{x})$.

As mentioned in Section 1, one of the most important approaches in optimization theory is to turn the nonlinear problems into linear problems, and then apply the usual techniques of linear programming. These linearizations, which are done approximately, are useful when the optimal solutions of two (nonlinear and linear) problems are the same. Assume that $x_0 \in \Omega$ is given. For each $\xi_0 \in \partial^\flat \vartheta(x_0)$ and $\xi_j \in \partial^\flat \psi_j(x_0)$, as $j \in J$, we define the linear functions $\varphi_{x_0}^{\xi_0} : \mathbb{R}^n \to \mathbb{R}$ and $\varphi_{x_0}^{\xi_j} : \mathbb{R}^n \to \mathbb{R}$ by

$$\varphi_{x_0}^{\xi_0}(x) := \langle \xi_0, x - x_0 \rangle, \quad \text{and} \quad \varphi_{x_0}^{\xi_j}(x) := \langle \xi_j, x - x_0 \rangle.$$

Also, we consider the following two linear optimization problems:

$$(\text{LP1}_{x_0}^{\xi_0}): \qquad \begin{aligned} &\inf \ \Big(\varphi_{x_0}^{\xi_0}(x) + \vartheta(x_0)\Big) \\ &\text{s.t.} \ \ \varphi_{x_0}^{\xi_j}(x) + \psi_j(x_0) \leq 0, \quad j \in J, \ \xi_j \in \partial^\flat \psi_j(x_0), \\ &\qquad x \in \mathbb{R}^n, \end{aligned}$$

$$(\text{LP2}_{x_0}^{\xi_0}): \qquad \begin{aligned} &\inf \ \Big(\varphi_{x_0}^{\xi_0}(x) + \vartheta(x_0)\Big) \\ &\text{s.t.} \ \ \varphi_{x_0}^{\xi_j}(x) + \psi_j(x_0) \leq 0, \quad j \in J(x_0), \ \xi_j \in \partial^\flat \psi_j(x_0), \\ &\qquad x \in \mathbb{R}^n. \end{aligned}$$

The feasible sets of the mentioned problems are denoted by $\Gamma 1_{x_0}$ and $\Gamma 2_{x_0}$, respectively, i.e.,

$$\Gamma 1_{x_0} := \left\{ x \in \mathbb{R}^n \mid \varphi_{x_0}^{\xi_j}(x) + \psi_j(x_0) \leq 0, \quad \forall j \in J, \forall \xi_j \in \partial^\flat \psi_j(x_0) \right\},$$

$$\Gamma 2_{x_0} := \left\{ x \in \mathbb{R}^n \mid \varphi_{x_0}^{\xi_j}(x) + \psi_j(x_0) \leq 0, \quad \forall j \in J(x_0), \forall \xi_j \in \partial^\flat \psi_j(x_0) \right\}.$$

Owing to $x_0 \in \Omega$ and $\langle \xi_j, x_0 - x_0 \rangle = 0$, we deduce that

$$\varphi_{x_0}^{\xi_j}(x_0) + \psi_j(x_0) = \psi_j(x_0) \leq 0, \qquad \forall j \in J.$$

The latter inequality and the definitions of $\Gamma 1_{x_0}$ and $\Gamma 2_{x_0}$ imply that

$$x_0 \in \Gamma 1_{x_0} \subseteq \Gamma 2_{x_0}, \qquad \forall x_0 \in \Omega. \tag{1}$$

Since the number of constraints of $(\mathrm{LP}1_{x_0}^{\xi_0})$ and $(\mathrm{LP}2_{x_0}^{\xi_0})$ are not finite (in general), they form a *linear semi-infinite programming problem* (LSIP). To study the properties of such problems, one can refer to [10]. The following two theorems show that, under some suitable assumptions, an optimal solution x_0 of (P) is also an optimal solution for $(\mathrm{LP}1_{x_0}^{\xi_0})$ and $(\mathrm{LP}2_{x_0}^{\xi_0})$.

Theorem 5. Suppose that $\hat{x} \in \Omega$ is an optimal solution of (P), and all the hypotheses of Theorem 3 hold. Then, $\hat{x}$ is an optimal solution of $(LP2_{\hat{x}}^{\hat{\xi}})$ for some $\hat{\xi} \in \partial^\flat \vartheta(\hat{x})$.

Proof. Employing Theorem 3, we can find some $\hat{\xi} \in \partial^\flat \vartheta(\hat{x})$, $\lambda_j \geq 0$ and $\xi_j \in \partial^\flat \psi_j(\hat{x})$ for $j \in J(\hat{x})$, such that

$$\hat{\xi} + \sum_{j \in J(\hat{x})} \lambda_j \xi_j = 0_n \implies \hat{\xi} = -\sum_{j \in J(\hat{x})} \lambda_j \xi_j.$$

Thus, for each $\bar{x} \in \Gamma 2_{\hat{x}}$, we get

$$\begin{aligned}
\varphi_{\hat{x}}^{\hat{\xi}}(\bar{x}) + \vartheta(\hat{x}) &= \langle \hat{\xi}, \bar{x} - \hat{x} \rangle + \vartheta(\hat{x}) = -\Big\langle \sum_{j \in J(\hat{x})} \lambda_j \xi_j, \bar{x} - \hat{x} \Big\rangle + \vartheta(\hat{x}) \\
&= -\sum_{j \in J(\hat{x})} \Big(\lambda_j \langle \xi_j, \bar{x} - \hat{x} \rangle \Big) + \vartheta(\hat{x}) = -\sum_{j \in J(\hat{x})} \lambda_j \overbrace{\varphi_{\hat{x}}^{\xi_j}(\bar{x}) + \vartheta(\hat{x})}^{\leq 0} \\
&\geq \vartheta(\hat{x}) = \langle \hat{\xi}, \hat{x} - \hat{x} \rangle + \vartheta(\hat{x}) = \varphi_{\hat{x}}^{\hat{\xi}}(\hat{x}) + \vartheta(\hat{x}).
\end{aligned}$$

Since $\bar{x}$ is an arbitrary element in $\Gamma 2_{\hat{x}}$, the above inequality means that $\hat{x}$ is an optimal solution of $(LP2_{\hat{x}}^{\hat{\xi}})$, and the proof is complete. □

Theorem 6. Suppose that $\hat{x} \in \Omega$ is an optimal solution of (P), and all the hypotheses of Theorem 3 hold. Then, $\hat{x}$ is an optimal solution of $(LP1_{\hat{x}}^{\hat{\xi}})$ for some $\hat{\xi} \in \partial^\flat \vartheta(\hat{x})$.

Proof. According to (1) and Theorem 5, the result is clear. □

The following theorem shows that the converse of Theorem 5 holds.

Theorem 7. Suppose that $\hat{x} \in \Omega$ is an optimal solution of $(LP2_{\hat{x}}^{\hat{\xi}})$ for some nonzero $\hat{\xi} \in \partial^{\flat}\vartheta(\hat{x}) \setminus \{0_n\}$. If ϑ is upper-semicontinuous at $\hat{x}$, then $\hat{x}$ is an optimal solution of (P).

Proof. Suppose on the contrary that there exists some $x^* \in \Omega$ such that

$$\vartheta(x^*) < \vartheta(\hat{x}). \tag{2}$$

Employing Theorem 4, we conclude that

$$\langle \hat{\xi}, x^* - \hat{x} \rangle \leq 0. \tag{3}$$

If $\left\langle \hat{\xi}, x^* - \hat{x} \right\rangle = 0$, we can find a sequence $\{w_k\}_{k=1}^{\infty}$ converging to $x^* - \hat{x}$ with $\left\langle \hat{\xi}, w_k \right\rangle > 0$, for all $k \in \mathbb{N}$. This inequality and Theorem 4 imply that

$$\left\langle \hat{\xi}, (w_k + \hat{x}) - \hat{x} \right\rangle > 0 \implies \vartheta(w_k + \hat{x}) \geq \vartheta(\hat{x}), \quad \forall k \in \mathbb{N}.$$

Thus, the upper-semicontinuity of ϑ allows us to deduce that

$$\vartheta(x^*) = \vartheta\Big(\lim_{k\to\infty}(w_k + \hat{x})\Big) \geq \lim_{k\to\infty} \vartheta(w_k + \hat{x}) \geq \vartheta(\hat{x}),$$

which contradicts (2). Therefore, $\left\langle \hat{\xi}, x^* - \hat{x} \right\rangle \neq 0$, and so

$$\left\langle \hat{\xi}, x^* - \hat{x} \right\rangle < 0, \tag{4}$$

by (3). On the other hand, owing to $x^* \in \Omega$, for each $j \in J(\hat{x})$, we have

$$\psi_j(x^*) \leq 0 = \psi_j(\hat{x}), \tag{5}$$

and by Theorem 4 we deduce for all $\xi_j \in \partial^{\flat}\psi_j(\hat{x})$ that

$$\varphi_{\hat{x}}^{\xi_j}(x^*) + \psi_j(\hat{x}) = \langle \xi_j, x^* - \hat{x} \rangle + 0 \leq 0.$$

Consequently, $x^* \in \Gamma 2_{\hat{x}}$, and by (4),

$$\varphi_{\hat{x}}^{\hat{\xi}}(x^*) + \vartheta(\hat{x}) = \langle \hat{\xi}, x^* - \hat{x} \rangle + \vartheta(\hat{x}) < \vartheta(\hat{x}) = \langle \hat{\xi}, \hat{x} - \hat{x} \rangle + \vartheta(\hat{x}) = \varphi_{\hat{x}}^{\hat{\xi}}(\hat{x}) + \vartheta(\hat{x}).$$

The above inequality contradicts the optimality of $\hat{x}$ for $(\text{LP}2_{\hat{x}}^{\hat{\xi}})$, and completes the proof. □

It is worth observing that we cannot repeat the proof of Theorem 7 for $(\text{LP}1_{\hat{x}}^{\hat{\xi}})$. In fact, in the proof of Theorem 7, the equality $\psi_j(\hat{x}) = 0$ is not true for all $j \in J$, and we cannot conclude that $x^* \in \Gamma 1_{\hat{x}}$. Consequently, the converse of Theorem 6 is not (in general) true, and we cannot replace $(\text{LP}2_{\hat{x}}^{\hat{\xi}})$ by $(\text{LP}1_{\hat{x}}^{\hat{\xi}})$ in Theorem 7. The following example shows the contents.

Example 1. Consider the following problem:

$$
\begin{aligned}
(Q): \qquad\qquad & \inf \ \ |x| \\
& \text{s.t.} \ \ x^3 - x \leq 0, \\
& \qquad x - 1 \leq 0, \\
& \qquad x \in \mathbb{R}.
\end{aligned}
$$

This problem has the form of (P), with $\vartheta(x) = |x|$, $J = \{1,2\}$, $\psi_1(x) = x^3 - x$, and $\psi_2(x) = x - 1$. Obviously, $\hat{x} := 0 \in \Omega$ with $J(0) = \{1\}$, and

$$
\partial^\flat \vartheta(0) = [-1,1], \qquad \partial^\flat \psi_1(0) = \{-1\}, \qquad \partial^\flat \psi_2(0) = \{1\}.
$$

Thus,

$$
\begin{aligned}
&\varphi_0^\xi(x) + \vartheta(0) = \xi(x-0) + 0 = \xi x, \qquad \text{for } \xi \in [-1,1], \\
&\varphi_0^{-1}(x) + \psi_1(0) = -1(x-0) + 0 = -x, \\
&\varphi_0^1(x) + \psi_2(0) = 1(x-0) - 1 = x - 1.
\end{aligned}
$$

Hence, for each $\xi \in [-1,1]$, the linearized problems (LP1_0^ξ) and (LP2_0^ξ) have the following forms:

$$
\begin{aligned}
(Q1_0^\xi): \qquad\qquad & \inf \ \ \xi x \\
& \text{s.t.} \ \ -x \leq 0, \\
& \qquad x - 1 \leq 0, \\
& \qquad x \in \mathbb{R},
\end{aligned}
$$

$$
\begin{aligned}
(Q2_0^\xi): \qquad\qquad & \inf \ \ \xi x \\
& \text{s.t.} \ \ -x \leq 0, \\
& \qquad x \in \mathbb{R}.
\end{aligned}
$$

These problems have the following feasible sets, respectively:

$$
\Gamma 1_0 = [0,1], \qquad \text{and} \qquad \Gamma 2_0 = [0,+\infty).
$$

Thus, we obtain the following assertions:

- $(Q2_0^\xi)$ has an optimal solution at $\hat{x}$ when $\xi > 0$, and Theorem 7 concludes that $\hat{x}$ is the optimal solution of (Q). This fact can be obtained trivially, by

$$
\begin{aligned}
\Omega = \Big\{x \in \mathbb{R} \mid x(x-1)(x+1) \leq 0,\ x - 1 \leq 0\Big\} = \\
\Big((-\infty,-1] \cup [0,1]\Big) \cap (-\infty,1] = (-\infty,-1] \cup [0,1].
\end{aligned}
$$

- $(Q1_0^\xi)$ has an optimal solution at $\hat{x}$ when $\xi > 0$, and $\hat{x}$ is the optimal solution of (Q). Thus, the converse of Theorem 6 can be true.

- $(Q1_0^{\xi})$ has an optimal solution at $x_1 := 1$ when $\xi < 0$, while x_1 is not an optimal solution of (Q).

Now, we consider the important concept of gap function for (P). A brief history and a detailed list of the advantages and properties of gap functions can be found in the pioneering paper [4]. One of the basic properties of gap function for optimization problems is its ability in characterizing the solutions of the problem in question. In order to do this for (P), we first formulate an appropriate gap function for the problem. We define a gap function Υ as follows:

$$\Upsilon : \bigcup_{x\in\Omega}\Big(\{x\}\times\partial^{\flat}\vartheta(x)\Big)\longrightarrow\mathbb{R},$$

$$\Upsilon(x,\xi) := \sup_{y\in\Omega}\langle\xi, x-y\rangle, \qquad \forall x\in\Omega,\ \forall\xi\in\partial^{\flat}\vartheta(x).$$

Equivalently, for each $x\in\Omega$ and $\xi\in\partial^{\flat}\vartheta(x)$, we have

$$\Upsilon(x,\xi) = \sup_{y}\Big\{\langle\xi, x-y\rangle \mid \vartheta_j(y)\le 0,\quad j\in J\Big\}.$$

According to $x\in\Omega$, we conclude that

$$\sup_{y}\Big\{\langle\xi, x-y\rangle \mid \vartheta_j(y)\le 0,\quad j\in J\Big\} \ge \langle\xi, x-x\rangle = 0,$$

and hence

$$\Upsilon(x,\xi)\ge 0, \qquad \forall x\in\Omega,\ \forall\xi\in\partial^{\flat}\vartheta(x). \tag{6}$$

We are going to show that the optimal solutions for (P) are characterized by the zero value of Υ.

Theorem 8. Suppose that $\Upsilon(\hat{x},\hat{\xi}) = 0$ for some $\hat{x}\in\Omega$ and for some nonzero $\hat{\xi}\in\partial^{\flat}\vartheta(\hat{x})\setminus\{0_n\}$. If ϑ is upper-semicontinuous at $\hat{x}$, then $\hat{x}$ is an optimal solution for (P).

Proof. Suppose on the contrary that there exists $y^*\in\Omega$ such that

$$\vartheta(y^*) < \vartheta(\hat{x}). \tag{7}$$

Similar to the proof of (4) we get

$$\langle\hat{\xi}, y^*-\hat{x}\rangle < 0. \tag{8}$$

So, by $y^*\in\Omega$ we deduce that

$$\begin{aligned}\Upsilon(\hat{x},\hat{\xi}) &= \sup_{y}\Big\{\langle\xi, x-y\rangle \mid \vartheta_j(y)\le 0,\quad j\in J\Big\}\\ &\ge \langle\hat{\xi}, \hat{x}-y^*\rangle = -\langle\hat{\xi}, y^*-\hat{x}\rangle > 0,\end{aligned}$$

where the last inequality holds by (8). The above inequality contradicts the assumption of $\Upsilon(\hat{x},\hat{\xi}) = 0$. So, (7) does not hold and the proof is complete. □

The following theorem presents the converse of Theorem 8.

Theorem 9. Suppose that $\hat{x} \in \Omega$ is an optimal solution for (P). Under the hypotheses of Theorem 3, $\Upsilon(\hat{x}, \hat{\xi}) = 0$ for some $\hat{\xi} \in \partial^{\flat} \vartheta(\hat{x})$.

Proof. According to Theorem 3, the equality

$$-\hat{\xi} = \sum_{j \in J(\hat{x})} \lambda_j \xi_j, \tag{9}$$

holds for some $\hat{\xi} \in \partial^{\flat} \vartheta(\hat{x})$, some $\xi_j \in \partial^{\flat} \psi_j(\hat{x})$ with $j \in J(\hat{x})$, and some $\lambda_j \geq 0$ with $j \in J(\hat{x})$. Assume that $y \in \Omega$ is chosen arbitrarily. Then, owing to Theorem 4, for each $j \in J(\hat{x})$ we have,

$$\psi_j(y) \leq 0 = \psi_j(\hat{x}) \Rightarrow \langle \xi_j, y - \hat{x} \rangle \leq 0.$$

The latter inequality and the fact that $\lambda_j \geq 0$ for $j \in J(\hat{x})$ imply

$$\sum_{j \in J(\hat{x})} \langle \lambda_j \xi_j, y - \hat{x} \rangle = \sum_{j \in J(\hat{x})} \lambda_j \langle \xi_j, y - \hat{x} \rangle \leq 0.$$

Combining this relation with (9), we obtain that

$$\langle \hat{\xi}, \hat{x} - y \rangle = \langle -\hat{\xi}, y - \hat{x} \rangle = \langle \sum_{j \in J(\hat{x})} \lambda_j \xi_j, y - \hat{x} \rangle \leq 0.$$

Since the above inequality holds for all $y \in \Omega$, we get

$$\Upsilon(\hat{x}, \hat{\xi}) = \sup_{y \in \Omega} \langle \hat{\xi}, \hat{x} - y \rangle \leq 0.$$

Due to the last inequality and (6), we deduce $\Upsilon(\hat{x}, \hat{\xi}) = 0$, as required. □

Theorem 10. Suppose that $\hat{x} \in \Omega$ is an optimal solution for (P). If ϑ is a convex function, then $\Upsilon(\hat{x}, \hat{\xi}) = 0$ for some $\hat{\xi} \in \partial \vartheta(\hat{x})$.

Proof. As the starting the proof, for each $j \in J$ set

$$\Omega_j := \{x \in \mathbb{R}^n \mid \psi_j(x) \leq 0\}.$$

Owing to the quasiconvexity of ψ_j for $j \in J$, we conclude that Ω_j is convex, and hence, Ω is convex by

$$\Omega = \bigcap_{j \in J} \Omega_j. \tag{10}$$

Since the convex function f attains its minimum on the convex set Ω at $\hat{x}$, Theorem 1 implies that

$$0_n \in \partial f(\hat{x}) + N(\Omega, \hat{x}).$$

This inclusion and the definition of convex normal cone imply that the following relation holds for some $\hat{\xi} \in \partial f(\hat{x})$:

$$-\hat{\xi} \in N(\Omega, \hat{x}) = \{z \in \mathbb{R}^n \mid \langle z, y - \hat{x} \rangle \le 0, \quad \forall y \in \Omega\}.$$

This means

$$\langle -\hat{\xi}, y - \hat{x} \rangle \le 0 \implies \langle \hat{\xi}, \hat{x} - y \rangle \le 0, \qquad y \in \Omega.$$

Therefore,

$$\Upsilon(\hat{x}, \hat{\xi}) = \sup_{y \in \Omega} \langle \hat{\xi}, \hat{x} - y \rangle \le 0.$$

The above inequality and (6) imply that $\Upsilon(\hat{x}, \hat{\xi}) = 0$, and the proof is complete. □

The following theorem shows that when $\vartheta(\cdot)$ and $\psi_j(\cdot)$ for $j \in J(\hat{x})$ are continuously differentiable at $\hat{x}$, the converse of Theorem 8 holds under the weakest constraint qualification GCQ, without any convexity assumption.

Theorem 11. Suppose that $\hat{x} \in \Omega$ is an optimal solution for (P), where ϑ and $\psi_j(\cdot)$ for $j \in J(\hat{x})$ are continuously differentiable at $\hat{x}$, and that GCQ holds at $\hat{x}$. Then, $\Upsilon\big(\hat{x}, \nabla\vartheta(\hat{x})\big) = 0$.

Proof. Employing Theorem 2, we find non-negative scalars λ_j for $j \in J$ such that

$$\nabla f(\hat{x}) + \sum_{J(\hat{x})} \lambda_j \nabla g_j(\hat{x}) = 0_n. \tag{11}$$

We claim that

$$\langle \nabla g_j(\hat{x}), z - \hat{x} \rangle \le 0, \qquad \forall j \in J(\hat{x}),\ \forall z \in \Omega. \tag{12}$$

Suppose, on the contrary, that $\langle g_{j^*}(\hat{x}), z^* - \hat{x} \rangle > 0$ for some $j^* \in J(\hat{x})$ and some $z^* \in \Omega$, i.e.,

$$0 < g'_{j^*}(\hat{x}; z^* - \hat{x}) = \lim_{\varepsilon \to 0^+} \frac{g_{j^*}\big(\hat{x} + \varepsilon(z^* - \hat{x})\big) - \overbrace{g_{j^*}(\hat{x})}^{=0}}{\varepsilon}.$$

This implies that $g_{j^*}\big(\hat{x} + \varepsilon(z^* - \hat{x})\big) > 0$, and so

$$\varepsilon z^* + (1 - \varepsilon)\hat{x} = \hat{x} + \varepsilon(z^* - \hat{x}) \notin \Omega.$$

Since $\hat{x} \in \Omega$ and $z^* \in \Omega$, the latter relation contradicts the convexity of Ω (by (10)), and this contradiction justifies the claim (12). Now, by (12), we have

$$\sum_{j \in J(\hat{x})} \langle \nabla \lambda_j g_j(\hat{x}), z - \hat{x} \rangle \le 0, \qquad \forall z \in \Omega. \tag{13}$$

On the other hand, equality (11) concludes

$$\langle \nabla f(\hat{x}), z - \hat{x} \rangle + \sum_{j \in J(\hat{x})} \langle \nabla \lambda_j g_j(\hat{x}), z - \hat{x} \rangle = 0, \qquad \forall z \in \Omega.$$

This equality together with (13) yields

$$\langle \nabla f(\hat{x}), z-\hat{x}\rangle \geq 0, \implies \langle \nabla f(\hat{x}), \hat{x}-z\rangle \leq 0, \qquad \forall z \in \Omega.$$

Therefore,

$$\Upsilon\Big(\hat{x}, \nabla f(\hat{x})\Big) = \sup_{z\in\Omega}\langle \nabla f(\hat{x}), \hat{x}-z\rangle \leq 0.$$

This inequality and (6) conclude $\Upsilon\Big(\hat{x}, \nabla f(\hat{x})\Big) = 0$, as required. □

Example 2. Consider the following problem:

$$\begin{array}{lll} (\mathcal{Y}): & \inf & x_1^2 + x_2^2 + x_1 - 1 \\ & \text{s.t.} & -x_1 \leq 0, \\ & & x_2 + 1 \leq 0, \\ & & x = (x_1, x_2) \in \mathbb{R}^2. \end{array}$$

Since $f(x) = x_1^2 + x_2^2 + x_1 - 1$, so $\nabla f(\hat{x}) = (2x_1 + 1, 2x_2)$. Considering

$$M = \{(x_1, x_2) \in \mathbb{R}^2 \mid -x_1 \leq 0,\ x_2 + 1 \leq 0\},$$

and the point $x^* = (0, -1) \in M$, we conclude that

$$\Upsilon\Big(x^*, \nabla f(x^*)\Big) = \sup_{(y_1,y_2)\in M} \langle (1,-2), (0-y_1, -1-y_2)\rangle = \sup_{(y_1,y_2)\in M} -y_1 + 2y_2 + 2.$$

The condition of $(y_1, y_2) \in M$ implies that $-y_1+2y_2+2 \leq 0$, and hence $\Upsilon\Big(x^*, \nabla f(x^*)\Big) = 0$. Consequently, x^* is an optimal solution of $(\mathcal{Y})$ by Theorem 8.
On the other hand, for each $\tilde{x} \in M$, we get

$$\begin{array}{rcl} \{(d_1, d_2) \in \mathbb{R}^2 \mid \langle (d_1, d_2), (-1.0)\rangle \leq 0,\ \langle (d_1, d_2), (0, 1)\rangle \leq 0\} & = & [0, +\infty) \times (-\infty, 0] \\ & = & conv\Big(\mathcal{Z}(M, \tilde{x})\Big), \end{array}$$

which implies that GCQ holds at all $\tilde{x} \in M$. Since, by a short calculation, one can see $\Upsilon\Big(\tilde{x}, \nabla f(\tilde{x})\Big) \neq 0$ for all $\tilde{x} \neq x^*$, Theorem 11 implies that $\tilde{x}$ is not an optimal solution of $(\mathcal{Y})$. Thus, the problem $(\mathcal{Y})$ has a unique optimal solution x^*.

4 Conclusion

In this paper, we derived two linearizations as well as a gap function for quasiconvex optimization problems with nonsmooth data. We expressed and proved our results in terms of incident subdifferential, using advanced quasiconvex optimization methods.

References

[1] Altangerel L., Bot R.I., Wanka G. (2006). "On gap functions for equilibrium problems via Fenchel duality", Pacific Journal of Optimization, 2, 667-678.

[2] Altangerel L., Bot R.I., Wanka G. (2007). “On the construction of gap functions for variational inequalities via conjugate duality”, Asia-Pacific Journal of Operational Research, 24, 353-371.

[3] Auslender A. (1976). “Optimisation: Méthods numériques”, Masson, Paris.

[4] Caristi G., Kanzi N., Soleimani-Damaneh M. (2018). “On gap functions for nonsmooth multi objective optimization problems”, Optimization Letters, 12, 273-286.

[5] Chen C.Y., Goh C.J., Yang X.Q. (1998). “The gap function of a convex multi criteria optimization problem”, European Journal of Operational Research, 111, 142-151.

[6] Giorgi G., Guerraggio A., Thierselder J. (2004). “Mathematics of optimization, smooth and nonsmooth cases”. Elsivier.

[7] Hassani Bafrani A., Sadeghieh A. (2018). “Quasi-gap and gap functions for non-smooth multi-objective semi-infinite optimization problems”, Control and Optimization in Applied Mathematics, 3, 1-12.

[8] Hearn DW. (1982). “The gap function of a convex program”, Operations Research Letters, 1, 67-71.

[9] Kanzi N., Sadeghieh A., Caristi G. (2019). “Optimality conditions for semi-infinite programming problems involving generalized convexity”, Optimization Letters, 13, 113-126.

[10] Kanzi N., Shaker Ardekani J., Caristi G. (2018). “Optimality, scalarization and duality in linear vector semi-infinite programming”, Optimization, 67, 523-536.

[11] Kanzi N., Soleymani-damaneh M. (2015). “Slater CQ, optimality and duality for quasiconvex semi-infinite optimization problems”, Journal of Mathematical Analysis and Applications, 434, 638-651.

[12] Lin M.H., Carlsson J.G., Ge D., Tsai J.F. (2013). “A review of piecewise linearization methods”, Mathematical problems in Engineering, 7, 14-25.

[13] López M.A., Vercher E. (1983). “Optimality conditions for nondifferentiable convex semi-infinite Programming”, Mathematical Programming, 27, 307-319.

[14] Penot J.P. (1998). “Are generalized derivatives useful for generalized convex functions? In generalized convexity, generalized monotonicity: Recent results”, J.P. Crouzeix, J. E. Martinez-Legaz, and M. Volle, (eds.), Kluwer, Dordrecht., 3-59.

[15] Penot J.P. (2000). “What is quasiconvex analysis?”, Optimization, 47, 35-110.

[16] Penot J.P., Zälinescu C. (2000). “Elements of quasiconûex subdifferential calculus”, Journal of Convex Analysis, 7, 243-269.

[17] Soleymani-damaneh M. (2008). “Infinite (semi-infinite) problems to characterize the optimality of nonlinear optimization problems”, European Journal of Operational Research, 188, 49-56.

[18] Soroush H. (2021). “Topological subdifferential and its role in nonsmooth optimization with quasi-convex data”, Control and Optimization in Applied Mathematics, 5, 83-91.

[19] Still C., Westerlund T. (2010). “A linear programming based optimization algorithm for solving nonlinear programming problems”, European Journal of Operational Research, 200, 658-670.

How to Cite this Article:
Soroush, H. (2022). "Linearization and gap function in nonsmooth quasiconvex optimization using incident subdifferential". Control and Optimization in Applied Mathematics, 7(1): 72-92. doi: 10.30473/COAM.2022.63088.1196