

# Optimization Letters

## Optimality conditions for pseudoconvex minimization over convex sets defined by tangentially convex constraints

--Manuscript Draft--

|  |  |
|--|--|
| <b>Manuscript Number:</b>                            | OPTL-D-14-00259R1  |
| <b>Full Title:</b>                                   | Optimality conditions for pseudoconvex minimization over convex sets defined by tangentially convex constraints  |
| <b>Article Type:</b>                                 | Manuscript   |
| <b>Keywords:</b>                                     | Convex Optimization, Nonsmooth Optimization, Optimality Conditions   |
| <b>Corresponding Author:</b>                         | Juan Enrique Martínez-Legaz<br>Universitat Autònoma de Barcelona<br>Bellaterra, Barcelona SPAIN  |
| <b>Corresponding Author Secondary Information:</b>   |  |
| <b>Corresponding Author's Institution:</b>           | Universitat Autònoma de Barcelona  |
| <b>Corresponding Author's Secondary Institution:</b> |  |
| <b>First Author:</b>                                 | Juan Enrique Martínez-Legaz  |
| <b>First Author Secondary Information:</b>           |  |
| <b>Order of Authors:</b>                             | Juan Enrique Martínez-Legaz  |
| <b>Order of Authors Secondary Information:</b>       |  |
| <b>Abstract:</b>                                     | We present necessary and sufficient optimality conditions for the minimization of pseudoconvex functions over convex sets defined by non necessarily convex functions, in terms of tangential subdifferentials. Our main result unifies a recent KKT type theorem obtained by Lasserre for differentiable functions with a nonsmooth version due to Dutta and Lalitha. |

# Optimality conditions for pseudoconvex minimization over convex sets defined by tangentially convex constraints

Juan Enrique Martínez-Legaz\*  
Departament d'Economia i d'Història Econòmica  
Universitat Autònoma de Barcelona  
08193 Bellaterra  
Spain

## Abstract

We present necessary and sufficient optimality conditions for the minimization of pseudoconvex functions over convex sets defined by non necessarily convex functions, in terms of tangential subdifferentials. Our main result unifies a recent KKT type theorem obtained by Lasserre for differentiable functions with a nonsmooth version due to Dutta and Lalitha.

## 1 Introduction

In the interesting article [3], Lasserre proved the surprising result that, in the case of differentiable problems satisfying the Slater constraint qualification and a mild nondegeneracy condition, for the KKT optimality conditions to be necessary and sufficient one can weaken the assumption that the constraint functions are convex, by just imposing the feasible set to be convex. A nonsmooth version of this result for locally Lipschitz functions has been recently obtained by Dutta and Lalitha [1], in terms of Clarke subdifferentials. This latter results uses the crucial assumption that the constraint functions are regular in the sense of Clarke, and therefore it does not generalize Lasserre result, since differentiable functions are not necessarily regular unless they are continuously differentiable. Thus the question arises whether it is possible to unify Lasserre's result with the one of Dutta and Lalitha, by obtaining a general statement of which those results would be corollaries. The present paper provides an affirmative answer to this question, by presenting a KKT type theorem (Theorem 9) based on the notion of tangential subdifferential, a notion due to Pshenichnyi [5]. Additionally,

---

\*This research was supported by the MICINN of Spain, Grant MTM2011-29064-C03-01, and under Australian Research Council's Discovery Projects funding scheme (project number DP140103213). The author is affiliated to MOVE (Markets, Organizations and Votes in Economics).

we show that the convexity assumption on the objective function, considered in [3] and [1], can be relaxed to pseudoconvexity, a condition which is not required for the necessity of the KKT conditions, but only for their sufficiency. This observation has also been made, in a differentiable context, by Giorgi in a recent paper [2].

Our method of proof essentially follows the same approach as the ones in [3] and [1], but our presentation is different and reveals the simple underlying geometry (see Proposition 1). We also give a more precise characterization of the convexity of the feasible set (Proposition 6), by showing that it is equivalent to its representability in a specific way by means of the directional derivatives of the constraint functions.

For convex analytic notions, we will use the same notation as in the classical book [6], and *rbd* we will denote relative boundary.

## 2 Results

We start with a simple geometric result, which will be fundamental in the characterization of convexity that underlies our main result.

**Proposition 1** *Let  $S \subseteq \mathbb{R}^n$  be such that  $ri S \neq \emptyset$  and  $C \subseteq \mathbb{R}^n$  be a convex set such that  $S \subseteq C \subseteq aff S$  and  $rbd S \subseteq rbd C$ , then  $C \subseteq cl ri S$ .*

**Proof.** Let  $x \in C$  and take  $\bar{x} \in ri S$ . Since  $aff S = aff C$ , we have  $ri S \subseteq ri C$ ; hence  $\bar{x} \in ri C$ . Therefore, for every  $\lambda \in [0, 1[$  one has  $(1 - \lambda)\bar{x} + \lambda x \in ri C$  and thus  $(1 - \lambda)\bar{x} + \lambda x \notin rbd C$ , which implies that  $(1 - \lambda)\bar{x} + \lambda x \notin rbd S$ . It follows that  $(1 - \lambda)\bar{x} + \lambda x \in ri S$  for every  $\lambda \in [0, 1[$ , since otherwise there would exist  $\bar{\lambda} \in ]0, 1[$  such that  $(1 - \bar{\lambda})\bar{x} + \bar{\lambda}x \notin S$  and hence, as  $\bar{x} \in ri S$ , by connectedness there would exist  $\lambda' \in ]0, \bar{\lambda}]$  such that  $(1 - \lambda')\bar{x} + \lambda'x \in rbd S$ , which is a contradiction. We conclude that  $x \in cl ri S$ . ■

**Remark 2** *It is easy to see that, under the assumptions  $ri S \neq \emptyset$  and  $rbd S \subseteq rbd C$ , the chain of inclusions  $S \subseteq C \subseteq aff S$  is equivalent to  $ri S \subseteq ri C$ .*

The next result is a straightforward consequence of Proposition 1.

**Corollary 3** *Let  $S \subseteq \mathbb{R}^n$  be a closed set such that  $ri S \neq \emptyset$ . The following statements are equivalent:*

- (i)  *$S$  is convex.*
- (ii)  *$rbd S = rbd conv S$ .*
- (iii)  *$rbd S \subseteq rbd conv S$ .*
- (iv)  *$conv S \subseteq cl ri S$ .*

**Proof.** Implications (i)  $\implies$  (ii)  $\implies$  (iii) are obvious. Implication (iii)  $\implies$  (iv) follows from Proposition 1 by taking  $C := conv S$ . Finally, if (iv) holds then  $conv S \subseteq cl ri S \subseteq cl S = S \subseteq conv S$ , which implies (i). ■

The following class of functions was introduced by Pshenichnyi [5]; they were called "tangentially convex" by Lemaréchal [4].

**Definition 4** A function  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  is called *tangentially convex* at  $x \in f^{-1}(\mathbb{R})$  if for every  $d \in \mathbb{R}^n$  the limit  $f'(x, d) := \lim_{t \rightarrow 0^+} \frac{f(x+td) - f(x)}{t}$  exists, is finite, and is a convex function of  $d$ .

It is worth noticing that, since  $f'(x, \cdot)$  is positively homogeneous, if  $f$  is tangentially convex at  $x$  then  $f'(x, \cdot)$  is sublinear.

The class of tangentially convex functions is rather large. It contains every convex function which has an open domain. Moreover, every function which is Gateaux differentiable at a point  $x$  is tangentially convex at  $x$ , since in such a case the directional derivative  $f'(x, \cdot)$  is linear. In particular, a function with open domain which is Gateaux differentiable everywhere is tangentially convex at each point of its domain, even though it is obviously not necessarily convex. Notice that the class of tangentially convex functions at a given point is a real vector space; hence, in particular, the sum of a convex function with a differentiable function provides an example of a tangentially convex function which, in general, is nonconvex and nondifferentiable. Another such example is the product of two nonnegative tangentially convex functions.

The notion of tangentially convex function suggests the introduction of an associated concept of subdifferential. The following definition is implicitly given in [5].

**Definition 5** The *tangential subdifferential* of  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  at  $x \in f^{-1}(\mathbb{R})$  is the set

$$\partial_T f(x) := \{x^* \in \mathbb{R}^n : \langle x^*, d \rangle \leq f'(x, d) \quad \forall d \in \mathbb{R}^n\}.$$

Clearly, if  $f$  is tangentially convex at  $x$ , from the sublinearity of  $f'(x, \cdot)$  it follows that  $\partial_T f(x) \neq \emptyset$  and  $f'(x, \cdot)$  is the support functional of  $\partial_T f(x)$ , that is, for every  $d \in \mathbb{R}^n$  one has

$$f'(x, d) = \max_{x^* \in \partial_T f(x)} \langle x^*, d \rangle.$$

In the particular case when  $f$  is convex, its tangential subdifferential at  $x$  coincides with the classical Fenchel subdifferential at  $x$ . In the case of a Gateaux differentiable function, its tangential subdifferential at a point is the singleton of the gradient. Tangential subdifferentials enjoy rich calculus rules. For instance, given two functions  $f$  and  $g$  which are tangentially convex at a common point  $x$ , one has  $\partial_T(f + g)(x) = \partial_T f(x) + \partial_T g(x)$ ; this additivity property easily follows from the relationship between tangential subdifferentials and directional derivatives, taking into account that the equality  $(f + g)'(x, d) = f'(x, d) + g'(x, d)$  holds for every  $d \in \mathbb{R}^n$ .

**Proposition 6** If the functions  $g_i : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  ( $i = 1, \dots, m$ ) are continuous, the system  $g_i(x) \leq 0$  ( $i = 1, \dots, m$ ) satisfies the Slater condition, and for every  $x \in S := \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \dots, m\}$  and  $i \in \{1, \dots, m\}$  such that  $g_i(x) = 0$  the function  $g_i$  is tangentially convex at  $x$  and  $\partial_T g_i(x) \neq \{0\}$ , then  $S$  is convex if and only if

$$S = \{y \in \mathbb{R}^n : g'_i(x, y - x) \leq 0 \quad \forall (x, i) \in S \times \{1, \dots, m\} \text{ s. t. } g_i(x) = 0\}. \quad (1)$$

**Proof.** Let  $C$  denote the right hand side of (1). This set is convex and closed, due to the tangential convexity assumption on the functions  $g_i$ ; hence, the "if" statement is obvious. Conversely, assume that  $S$  is convex. The inclusion  $\subseteq$  in (1) easily follows from the definition of directional derivative. To prove the opposite inclusion, in view of Proposition 1 we only need to prove that  $bd S \subseteq bd C$ . Let  $x \in bd S$ . We have  $x \in C$  and, since the functions  $g_i$  are continuous,  $g_i(x) = 0$  for some  $i \in \{1, \dots, m\}$ . Take  $x^* \in \partial_T g_i(x) \setminus \{0\}$ ; then  $\langle x^*, y - x \rangle \leq g'_i(x, y - x) \leq 0$  for every  $y \in C$ , which shows that the hyperplane  $\langle x^*, y - x \rangle = 0$  supports  $C$  at  $x$ . Therefore,  $x \in bd C$ . We have thus proved the required inclusion  $bd S \subseteq bd C$ . ■

We next extend the well known notion of pseudoconvexity for differentiable functions to the tangentially convex setting.

**Definition 7** A function  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  which is tangentially convex at  $x \in f^{-1}(\mathbb{R})$  is said to be pseudoconvex at  $x$  if  $f(y) \geq f(x)$  for every  $y \in \mathbb{R}^n$  such that  $f'(x, y - x) \geq 0$ .

Notice that, for a tangentially convex function, pseudoconvexity is a much weaker condition than convexity; for instance, every Gateaux differentiable pseudoconvex function is tangentially convex but not necessarily convex (consider, e.g., the one variable function  $f$  given by  $f(x) = -e^x$ ).

In the proof of the main theorem, we will use the following lemma.

**Lemma 8** Let  $s : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be a sublinear function. If  $s$  vanishes on an open set, then it is nonnegative everywhere.

**Proof.** Assume that  $s$  vanishes on the open ball  $B$  with center  $c$  and radius  $r$ , and let  $x \in \mathbb{R}^n$ . Set  $b := c - \frac{r}{2\|x\|}x$ ; since  $b \in B$ , by the sublinearity of  $s$  we have  $s(x) \geq \frac{2\|x\|}{r}(s(c) - s(b)) = 0$ . ■

The following theorem is the main result in this paper.

**Theorem 9** Let the functions  $g_i : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  ( $i = 1, \dots, m$ ) satisfy all the assumptions of Proposition 6. Assume further that the set

$$S := \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \dots, m\}$$

is convex, the function  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  is tangentially convex at  $\bar{x} \in S$ . If either  $g_i(\bar{x}) < 0$  for every  $i \in \{1, \dots, m\}$  or there exists some  $i \in \{1, \dots, m\}$  such that  $g_i(\bar{x}) = 0$  and  $0 \notin \partial_T g_i(\bar{x})$ , a necessary condition for  $\bar{x}$  to be a global minimizer of  $f$  over  $S$  is the existence of real numbers  $\lambda_i \geq 0$  ( $i = 1, \dots, m$ ) such that

$$i) 0 \in \partial_T f(\bar{x}) + \sum_{i=1}^m \lambda_i \partial_T g_i(\bar{x}),$$

$$ii) \lambda_i g_i(\bar{x}) = 0, \forall i = 1, \dots, m.$$

If  $f$  is pseudoconvex at  $\bar{x}$ , this condition is sufficient.

**Proof.** If  $\bar{x}$  is a global minimizer, then, by [5, p. 88, Corollary], there exist real numbers  $\lambda_i \geq 0$  ( $i = 0, 1, \dots, m$ ), not all zero, satisfying ii) and  $\lambda_0 f'(\bar{x}, d) + \sum_{i=1}^m \lambda_i g'_i(\bar{x}, d) \geq 0, \forall d \in \mathbb{R}^n$ . In particular, setting  $I := \{i \in \{1, \dots, m\} : \lambda_i > 0\}$ , for a Slater point  $\hat{x}$  we have

$$\begin{aligned} \lambda_0 f'(\bar{x}, \hat{x} - \bar{x}) + \sum_{i \in I} \lambda_i g'_i(\bar{x}, \hat{x} - \bar{x}) &= \lambda_0 f'(\bar{x}, \hat{x} - \bar{x}) + \sum_{i=1}^m \lambda_i g'_i(\bar{x}, \hat{x} - \bar{x}) \\ &\geq 0. \end{aligned}$$

Hence, if we had  $\lambda_0 = 0$ , then, by ii) and (1), we would have  $g'_i(\bar{x}, \hat{x} - \bar{x}) = 0$  for all  $i \in I$ ; moreover,  $I \neq \emptyset$ , which implies the existence of some  $i \in \{1, \dots, m\}$  such that  $g_i(\bar{x}) = 0$ . Since the functions  $g_i$  are continuous, the set of Slater points is open, and therefore, using Lemma 8, we can easily prove that  $g'_i(\bar{x}, d) \geq 0$  for all  $i \in I$  and  $d \in \mathbb{R}^n$ , that is,  $0 \in \partial_T g_i(\bar{x})$ , which is a contradiction. Thus  $\lambda_0 > 0$ , and without loss of generality we can take  $\lambda_0 = 1$ . Then, for every  $d \in \mathbb{R}^n$  we have

$$\max_{x^* \in \partial_T f(\bar{x}) + \sum_{i=1}^m \lambda_i \partial_T g_i(\bar{x})} \langle x^*, d \rangle = f'(\bar{x}, d) + \sum_{i=1}^m \lambda_i g'_i(\bar{x}, d) \geq 0,$$

which, since the set  $\partial_T f(\bar{x}) + \sum_{i=1}^m \lambda_i \partial_T g_i(\bar{x})$  is convex and closed, implies i).

Conversely, assume that  $f$  is pseudoconvex at  $\bar{x}$  and there exist real numbers  $\lambda_i \geq 0$  ( $i = 1, \dots, m$ ), not all zero, such that i) and ii) hold. Then, by i), for every  $x \in S$  we have

$$\begin{aligned} f'(\bar{x}, x - \bar{x}) + \sum_{i \in I} \lambda_i g'_i(\bar{x}, x - \bar{x}) &= \max_{x^* \in \partial_T f(\bar{x}) + \sum_{i \in I} \lambda_i \partial_T g_i(\bar{x})} \langle x^*, x - \bar{x} \rangle \\ &\geq 0. \end{aligned}$$

Hence, as from ii) and Proposition 6 it follows that  $g'_i(\bar{x}, x - \bar{x}) \leq 0$  for every  $i \in I$ , we deduce that  $f'(\bar{x}, x - \bar{x}) \geq 0$ ; therefore, by the pseudoconvexity of  $f$  at  $\bar{x}$ , we have  $f(x) \geq f(\bar{x})$ . This proves that  $\bar{x}$  is a global minimizer of  $f$  over  $S$ . ■

To conclude, we show that the KKT type theorems due to Lasserre and to Dutta and Lalitha easily follow from Theorem 9.

**Corollary 10** [3, Thm. 2.3] *Let the functions  $f, g_i : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  ( $i = 1, \dots, m$ ) be differentiable, the set  $S := \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \dots, m\}$  be convex, and assume that the system  $g_i(x) \leq 0$  ( $i = 1, \dots, m$ ) satisfies the Slater condition and the nondegeneracy property*

$$\nabla g_i(x) \neq 0 \text{ for every } x \in S \text{ and } i \in \{1, \dots, m\} \text{ such that } g_i(x) = 0.$$

If  $f$  is convex then, a necessary and sufficient condition for  $\bar{x} \in S$  to be a global minimum of  $f$  over  $S$  is the existence of real numbers  $\lambda_i \geq 0$  ( $i = 1, \dots, m$ ) such that

$$\begin{aligned} i) \quad & \nabla f(\bar{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}) = 0, \\ ii) \quad & \lambda_i g_i(\bar{x}) = 0, \forall i = 1, \dots, m. \end{aligned}$$

**Proof.** It is an immediate consequence of Theorem 9, since a differentiable function  $f$  is tangentially convex at every point  $x$ , with  $\partial_T f(x) = \{\nabla f(x)\}$ , and every convex function is pseudoconvex at every point. ■

**Corollary 11** [1, Thm. 2.4] *Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be convex, the functions  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$  ( $i = 1, \dots, m$ ) be locally Lipschitz and regular in the sense of Clarke, the set  $S := \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \dots, m\}$  be convex, and assume that the system  $g_i(x) \leq 0$  ( $i = 1, \dots, m$ ) satisfies the Slater condition and the nondegeneracy property*

$$0 \notin \partial^\circ g_i(x) \text{ for every } x \in S \text{ and } i \in \{1, \dots, m\} \text{ such that } g_i(x) = 0.$$

*Then a necessary and sufficient condition for  $\bar{x} \in S$  to be a global minimum of  $f$  over  $S$  is the existence of real numbers  $\lambda_i \geq 0$  ( $i = 1, \dots, m$ ) such that*

$$\begin{aligned} i) \quad & 0 \in \partial f(\bar{x}) + \sum_{i=1}^m \lambda_i \partial^\circ g_i(\bar{x}) = 0, \\ ii) \quad & \lambda_i g_i(\bar{x}) = 0, \forall i = 1, \dots, m. \end{aligned}$$

**Proof.** It is an immediate consequence of Theorem 9, since every locally Lipschitz regular function  $g$  is tangentially convex at every point  $x$ , with  $\partial_T g(x) = \partial^\circ g(x)$ , and a convex function  $f$  is pseudoconvex and tangentially convex at every point  $x$ , with  $\partial_T f(x) = \partial f(x)$ . ■

**Acknowledgment.** I am grateful to two anonymous referees for their useful comments, which have helped me improve the presentation.

## References

- [1] Dutta, J.; Lalitha, C. S.: Optimality conditions in convex optimization revisited. *Optim. Lett.* 7 (2013), no. 2, 221–229.
- [2] Giorgi, G.: Optimality conditions under generalized convexity revisited. *Ann. Univ. Buchar. Math. Ser. 4(LXII)* (2013), no. 2, 479–490.
- [3] Lasserre, J. B.: On representations of the feasible set in convex optimization. *Optim. Lett.* 4 (2010), no. 1, 1–5.
- [4] Lemaréchal, C.: An introduction to the theory of nonsmooth optimization. *Optimization* 17 (1986), no. 6, 827–858.

- [5] Pshenichnyi, B. N.: Necessary conditions for an extremum. Marcel Dekker, Inc., New York 1971.
- [6] Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton 1970.