



HAL
open science

On representations of the feasible set in convex optimization

Jean-Bernard Lasserre

► **To cite this version:**

Jean-Bernard Lasserre. On representations of the feasible set in convex optimization. Optimization Letters, 2010, 4 (1), p. 1-5. <10.1007/s11590-009-0153-6>. <hal-00430141>

HAL Id: hal-00430141

<https://hal.science/hal-00430141v1>

Submitted on 5 Nov 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

ON REPRESENTATIONS OF THE FEASIBLE SET IN CONVEX OPTIMIZATION

JEAN B. LASSERRE

ABSTRACT. We consider the convex optimization problem $\min_{\mathbf{x}} \{f(\mathbf{x}) : g_j(\mathbf{x}) \leq 0, j = 1, \dots, m\}$ where f is convex, the feasible set \mathbf{K} is convex and Slater's condition holds, but the functions g_j 's are *not* necessarily convex. We show that for any representation of \mathbf{K} that satisfies a mild nondegeneracy assumption, every minimizer is a Karush-Kuhn-Tucker (KKT) point and conversely every KKT point is a minimizer. That is, the KKT optimality conditions are necessary and sufficient as in convex programming where one assumes that the g_j 's are convex. So in convex optimization, and as far as one is concerned with KKT points, what really matters is the geometry of \mathbf{K} and not so much its representation.

1. INTRODUCTION

Given differentiable functions $f, g_j : \mathbb{R}^n \rightarrow \mathbb{R}, j = 1, \dots, m$, consider the following convex optimization problem:

$$(1.1) \quad f^* := \inf_{\mathbf{x}} \{ f(\mathbf{x}) : \mathbf{x} \in \mathbf{K} \}$$

where f is convex and the feasible set $\mathbf{K} \subset \mathbb{R}^n$ is convex and represented in the form:

$$(1.2) \quad \mathbf{K} = \{ \mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \leq 0, j = 1, \dots, m \}.$$

Convex optimization usually refers to minimizing a convex function over a convex set without precisising its representation (see e.g. Ben-Tal and Nemirovsky [1, Definition 5.1.1] or Bertsekas et al. [3, Chapter 2]), and it is well-known that convexity of the function f and of the set \mathbf{K} imply that every local minimum is a global minimum. An elementary proof only uses the geometry of \mathbf{K} , not its representation by the defining functions g_j ; see e.g. Bertsekas et al. [3, Prop. 2.1.2].

The convex set \mathbf{K} may be represented by different choices of the (not necessarily convex) defining functions $g_j, j = 1, \dots, m$. For instance, the set

$$\mathbf{K} := \{ \mathbf{x} \in \mathbb{R}^2 : 1 - x_1x_2 \leq 0; \mathbf{x} \geq 0 \}$$

is convex but the function $\mathbf{x} \mapsto 1 - x_1x_2$ is not convex on \mathbb{R}_+^2 . Of course, depending on the choice of the defining functions (g_j), several properties may or may not hold. In particular, the celebrated Karush-Kuhn-Tucker (KKT) optimality conditions *depend* on the representation of \mathbf{K} . Recall that $\mathbf{x} \in \mathbf{K}$ is a KKT point if

$$(1.3) \quad \nabla f(\mathbf{x}) + \sum_{j=1}^m \lambda_j \nabla g_j(\mathbf{x}) = 0 \quad \text{and} \quad \lambda_j g_j(\mathbf{x}) = 0, j = 1, \dots, m,$$

for some nonnegative vector $\lambda \in \mathbb{R}^m$. (More precisely (\mathbf{x}, λ) is a KKT point.)

1991 *Mathematics Subject Classification.* 90C25 65K05.

Key words and phrases. Convex programming; KKT conditions.

Convex programming refers to the situation where f is convex *and* the defining functions g_j of \mathbf{K} are also convex. See for instance Ben-Tal and Nemirovsky [1, p. 335], Berkovitz [2, p. 179], Boyd and Vandenberghe [4, p. 7], Bertsekas et al. [3, §3.5.5], Nesterov and Nemirovskii [6, p. 217-218], and Hiriart-Urruty [5].

A crucial feature of convex programming is that when Slater's condition holds¹, the KKT optimality conditions (1.3) are necessary and sufficient, which shows that a representation of the convex set \mathbf{K} with convex functions (g_j) has some very attractive features.

The purpose of this note is to show that in fact, when \mathbf{K} is convex and as far as one is concerned with KKT points, what really matters is the geometry of \mathbf{K} and not so much its representation. Indeed, we show that if \mathbf{K} is convex and Slater's condition holds then the KKT optimality conditions (1.3) are also necessary and sufficient for *all* representations of \mathbf{K} that satisfy a mild nondegeneracy condition, no matter if the g_j 's are convex. So this attractive feature is not specific to representations of \mathbf{K} with convex functions.

That a KKT point is a local (hence global) minimizer follows easily from the convexity of \mathbf{K} . More delicate is the fact that any local (hence global) minimizer is a KKT point. Various *constraint qualifications* are usually required to hold at a minimizer, and when the g_j 's are convex the simple Slater's condition is enough. Here we show that Slater's condition is also sufficient for all representations of \mathbf{K} that satisfy a mild additional nondegeneracy assumption on the boundary of \mathbf{K} . Moreover under Slater's condition this mild nondegeneracy assumption is automatically satisfied if the g_j 's are convex.

2. MAIN RESULT

Let $\mathbf{K} \subset \mathbb{R}^n$ be as in (1.2). We first start with the following non degeneracy assumption:

Assumption 2.1 (nondegeneracy). *For every $j = 1, \dots, m$,*

$$(2.1) \quad \nabla g_j(\mathbf{x}) \neq 0, \quad \text{whenever } \mathbf{x} \in \mathbf{K} \text{ and } g_j(\mathbf{x}) = 0.$$

Observe that under Slater's condition, (2.1) is automatically satisfied if g_j is convex. Indeed if $g_j(\mathbf{x}) = 0$ and $\nabla g_j(\mathbf{x}) = 0$ then by convexity 0 is the global minimum of g_j on \mathbb{R}^n . Hence there is no $\mathbf{x}_0 \in \mathbf{K}$ with $g_j(\mathbf{x}_0) < 0$. We next state the following characterization of convexity.

Lemma 2.2. *With $\mathbf{K} \subset \mathbb{R}^n$ as in (1.2), let Assumption 2.1 and Slater's condition both hold for \mathbf{K} . Then \mathbf{K} is convex if and only if for every $j = 1, \dots, m$:*

$$(2.2) \quad \langle \nabla g_j(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \leq 0, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbf{K} \text{ with } g_j(\mathbf{x}) = 0.$$

Proof. Only if part. Assume that \mathbf{K} is convex and $\langle \nabla g_j(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle > 0$ for some $j \in \{1, \dots, m\}$ and some $\mathbf{x}, \mathbf{y} \in \mathbf{K}$ with $g_j(\mathbf{x}) = 0$. Then $g_j(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) > 0$ for all sufficiently small t , in contradiction with $\mathbf{x} + t(\mathbf{y} - \mathbf{x}) \in \mathbf{K}$ for all $0 \leq t \leq 1$ (by convexity of \mathbf{K}).

If part. By (2.2), at every point \mathbf{x} on the boundary of \mathbf{K} , there exists a supporting hyperplane for \mathbf{K} . As \mathbf{K} is closed with nonempty interior, by [8][Th. 1.3.3] the set \mathbf{K} is convex². \square

¹Slater's condition holds for \mathbf{K} if for some $\mathbf{x}_0 \in \mathbf{K}$, $g_j(\mathbf{x}_0) < 0$ for every $j = 1, \dots, m$.

²The author wishes to thank Prof. L. Tunçel for providing him with the reference [8].

Theorem 2.3. *Consider the nonlinear programming problem (1.1) and let Assumption 2.1 and Slater's condition both hold. If f is convex then every minimizer is a KKT point and conversely, every KKT point is a minimizer.*

Proof. Let $\mathbf{x}^* \in \mathbf{K}$ be a minimizer (hence a global minimizer) with $f^* = f(\mathbf{x}^*)$. We first prove that \mathbf{x}^* is a KKT point. The Fritz-John optimality conditions state that

$$\lambda_0 \nabla f(\mathbf{x}^*) + \sum_{j=1}^m \lambda_j \nabla g_j(\mathbf{x}^*) = 0; \quad \lambda_j g_j(\mathbf{x}^*) = 0, \quad j = 1, \dots, m,$$

for some non trivial nonnegative vector $0 \neq \lambda \in \mathbb{R}^{m+1}$. See e.g. Hiriart-Urruty [5, Th. page 77] or Polyak [7, Theor. 1, p. 271]. We next prove that $\lambda_0 \neq 0$. Suppose that $\lambda_0 = 0$ and let $J := \{j \in \{1, \dots, m\} : \lambda_j > 0\}$. As $\lambda \neq 0$ and $\lambda_0 = 0$, the set J is nonempty. Next, as $g_j(\mathbf{x}_0) < 0$ for every $j = 1, \dots, m$, there is some $\rho > 0$ such that $B(\mathbf{x}_0, \rho) := \{\mathbf{z} \in \mathbb{R}^n : \|\mathbf{z} - \mathbf{x}_0\| < \rho\} \subset \mathbf{K}$ and $g_j(\mathbf{z}) < 0$ for all $\mathbf{z} \in B(\mathbf{x}_0, \rho)$ and all $j \in J$. Therefore we obtain

$$\sum_{j \in J} \lambda_j \langle \nabla g_j(\mathbf{x}^*), \mathbf{z} - \mathbf{x}^* \rangle = 0 \quad \forall \mathbf{z} \in B(\mathbf{x}_0, \rho),$$

which, by Lemma 2.2, implies that $\langle \nabla g_j(\mathbf{x}^*), \mathbf{z} - \mathbf{x}^* \rangle = 0$ for every $j \in J$ and every $\mathbf{z} \in B(\mathbf{x}_0, \rho)$. But this clearly implies that $\nabla g_j(\mathbf{x}^*) = 0$ for every $j \in J$, in contradiction with Assumption 2.1. Hence $\lambda_0 > 0$ and we may and will set $\lambda_0 = 1$, so that the KKT conditions hold at \mathbf{x}^* .

Conversely, let $\mathbf{x} \in \mathbf{K}$ be an arbitrary KKT point, i.e., $\mathbf{x} \in \mathbf{K}$ satisfies

$$\nabla f(\mathbf{x}) + \sum_{j=1}^m \lambda_j \nabla g_j(\mathbf{x}) = 0; \quad \lambda_j g_j(\mathbf{x}) = 0, \quad j = 1, \dots, m,$$

for some nonnegative vector $\lambda \in \mathbb{R}^m$. Suppose that there exists $\mathbf{y} \in \mathbf{K}$ with $f(\mathbf{y}) < f(\mathbf{x})$. Then we obtain the contradiction:

$$\begin{aligned} 0 &> f(\mathbf{y}) - f(\mathbf{x}) \\ &\geq \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \quad [\text{by convexity of } f] \\ &= - \sum_{j=1}^m \lambda_j \langle \nabla g_j(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq 0 \end{aligned}$$

where the last inequality follows from $\lambda \geq 0$ and Lemma 2.2. Hence \mathbf{x} is a minimizer. \square

Hence if \mathbf{K} is convex and both Assumption 2.1 and Slater's condition hold, there is a one-to-one correspondence between KKT points and minimizers. That is, the KKT optimality conditions are necessary and sufficient for all representations of \mathbf{K} that satisfy Slater's condition and Assumption 2.1.

However there is an important additional property when all the defining functions g_j are convex. Dual methods of the type

$$\sup_{\lambda \in \mathbb{R}_+^m} \left\{ \inf_{\mathbf{x}} f(\mathbf{x}) + \sum_{j=1}^m \lambda_j g_j(\mathbf{x}) \right\},$$

are well defined because $\mathbf{x} \mapsto f(\mathbf{x}) + \sum_{j=1}^m \lambda_j g_j(\mathbf{x})$ is a convex function. In particular, the Lagrangian $\mathbf{x} \mapsto L_f(\mathbf{x}) := f(\mathbf{x}) - f^* + \sum_{j=1}^m \lambda_j g_j(\mathbf{x})$, defined from an

arbitrary KKT point $(\mathbf{x}^*, \lambda) \in \mathbf{K} \times \mathbb{R}_+^m$, is convex and nonnegative on \mathbb{R}^n , with \mathbf{x}^* being a global minimizer. If the g_j 's are not convex this is not true in general.

Example 1. Let $n = 2$ and consider the problem

$$\mathbf{P} : f^* = \min \{ f(\mathbf{x}) : a - x_1x_2 \leq 0; \mathbf{Ax} \leq \mathbf{b}; \mathbf{x} \geq 0 \},$$

where $a > 0$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, and f is convex and differentiable. The set

$$\mathbf{K} := \{ \mathbf{x} \in \mathbb{R}^2 : a - x_1x_2 \leq 0; \mathbf{Ax} \leq \mathbf{b}; \mathbf{x} \geq 0 \}$$

is convex and it is straightforward to check that Assumption 2.1 holds. Therefore, by Theorem 2.3, if Slater's condition holds, every KKT point is a global minimizer. However, the Lagrangian

$$\mathbf{x} \mapsto f(\mathbf{x}) - f^* + \psi(a - x_1x_2) + \langle \lambda, \mathbf{Ax} - \mathbf{b} \rangle - \langle \mu, \mathbf{x} \rangle,$$

with nonnegative $(\psi, \lambda, \mu) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^n$, may not be convex whenever $\psi \neq 0$ (for instance if f is linear). On the other hand, notice that \mathbf{K} has the equivalent convex representation

$$\mathbf{K} := \left\{ \mathbf{x} \in \mathbb{R}^2 : \begin{bmatrix} x_1 & \sqrt{a} \\ \sqrt{a} & x_2 \end{bmatrix} \succeq 0; \mathbf{Ax} \leq \mathbf{b} \right\},$$

where for a real symmetric matrix \mathbf{B} , the notation $\mathbf{B} \succeq 0$ stands for \mathbf{B} is positive semidefinite.

A topic of further investigation is concerned with computational efficiency. Can *efficient* algorithms be devised for some class of convex problems (1.1) where the defining functions g_j of \mathbf{K} are not necessarily convex?

REFERENCES

- [1] A. Ben-Tal, A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, SIAM, Philadelphia, 2001.
- [2] L.D. Berkovitz. *Convexity and Optimization in \mathbb{R}^n* , John Wiley & Sons, Inc., 2002.
- [3] D. Bertsekas, A. Nedić, E. Ozdaglar. *Convex Analysis and Optimization*, Athena Scientific, Belmont, Massachusetts, 2003.
- [4] S. Boyd, L. Vandenberghe. *Convex Optimization*, Cambridge University Press, Cambridge, 2004.
- [5] J.-B. Hiriart-Urruty. *Optimisation et Analyse Convexe*, Presses Universitaires de France, 1998.
- [6] Y. Nesterov, A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.
- [7] B.T. Polyak. *Introduction to Optimization*, Optimization Software, Inc., New York, 1987.
- [8] R. Schneider. *Convex Bodies: The Brunn-Minkowski Theory*, Cambridge University Press, Cambridge, UK (1994).

LAAS-CNRS AND INSTITUTE OF MATHEMATICS, UNIVERSITY OF TOULOUSE, LAAS, 7 AVENUE
DU COLONEL ROCHE, 31077 TOULOUSE CÉDEX 4, FRANCE
E-mail address: `lasserre@laas.fr`