A Simple Characterization of Solution Sets of Convex Programs

O. L. Mangasarian

The purpose of this work is to give some useful and simple properties of solution sets of convex programs. Surprisingly, these results have not been given before, to the best knowledge of the author. For example, given a differentiable convex function f on the n-dimensional real space R^n and given a convex set X in R^n, the solution set of the minimization problem of f on X is characterized by a constant gradient of the objective as well as a constant scalar product of the gradient with the point at which the gradient is evaluated. This result can be employed, for example, to show the polyhedrality of certain solution sets as well as to characterize their boundedness. For nonsmooth convex objective functions we show that the subdifferential of the objective function is constant on the relative interior of the solution set and give a number of useful characterizations of this set.

We begin with a lemma which shows that the gradient of the objective function of a convex programming problem is constant on the solution set of the problem.

1. Lemma Let X be a convex set in the real n-space R^n and let f be a twice continuously differentiable convex function on some open convex set containing X. Let $\bar{X}$ be the nonempty solution set of $\min\limits_{x \in X} f(x)$. Then the gradient $\triangledown f(x)$ is constant on $\bar{X}$.

Proof Let $x,\ y \in \bar{X}$. Then by the minimum principle [5, Theorem 9.3.3], [2, Page 121, Proposition 1.3]

$$\triangledown f(y)(y-x) \geq \triangledown f(x)(y-x) \geq 0$$

$$\triangledown f(x)(x-y) \geq \triangledown f(y)(x-y) \geq 0$$

Hence

$$\triangledown f(y)(y-x) = \triangledown f(x)(y-x) = 0 \tag{1}$$

Now

$$\begin{aligned}\triangledown f(y) - \triangledown f(x) &= \Big[\triangledown f\big(x+t(y-x)\big)\Big]_{t=0}^{t=1} = \int_{t=0}^{t=1} \triangledown^2 f\big(x+t(y-x)\big)(y-x)dt \\ &= C(y-x),\end{aligned} \tag{2}$$

where C is the $n \times n$ symmetric matrix defined by

$$C := \int_{t=0}^{t=1} \nabla^2 f\big(x + t(y-x)\big)dt$$

Since $zCz \geq 0$ for all z in R^n it follows that C is positive semidefinite. We have then from (1) and (2) that

$$0 = (y-x)\big(\nabla f(y) - \nabla f(x)\big) = (y-x)C(y-x) \tag{3}$$

Since C is symmetric positive semidefinite, it follows from (3) and (2) that

$$0 = C(y-x) = \nabla f(y) - \nabla f(x). \quad \blacksquare$$

The above lemma is extended in Lemma 1a in the sequel to a nondifferentiable convex f. In fact, by using Lemma 1a we can state Lemma 1 under the weaker assumption that f is convex and differentiable.

We now characterize the solution set of a convex program in terms of <u>any</u> of its solution points.

2. Theorem (Characterization of the solution set of a convex program) Let X be a convex set in R^n, let f be twice continuously differentiable convex function on some open convex set containing X and let $\bar{x}$ be any point in the solution set $\bar{X}$ of $\min\limits_{x \in X} f(x)$. Then $\bar{X} = \bar{S} = \hat{S}$, where

$$\begin{aligned} \bar{S} &:= \{x | x \in X,\ \nabla f(\bar{x})(x-\bar{x}) = 0,\ \nabla f(x) = \nabla f(\bar{x})\} \\ \hat{S} &:= \{x | x \in X,\ \nabla f(\bar{x})(x-\bar{x}) \leq 0,\ \nabla f(x) = \nabla f(\bar{x})\} \end{aligned} \tag{4}$$

Proof $(\hat{S} \subset \bar{X})$ Let $x \in \hat{S}$. It follows then from the convexity of f that

$$f(\bar{x}) - f(x) \geq \nabla f(x)(\bar{x}-x) = \nabla f(\bar{x})(\bar{x}-x) \geq 0$$

Hence $f(x) \leq f(\bar{x})$ and $x \in \bar{X}$. Thus $\hat{S} \subset \bar{X}$.

$(\bar{X} \subset \bar{S})$ Let $x \in \bar{X}$. By Lemma 1 it follows that $\nabla f(x) = \nabla f(\bar{x})$. By the minimum principle we have that

$$\nabla f(\bar{x})(x-\bar{x}) \geq 0 \tag{5}$$

$$\nabla f(\bar{x})(\bar{x}-x) \geq 0 \tag{6}$$

Thus $\nabla f(\bar{x})(x-\bar{x})=0$ and $x \in \bar{S}$. Hence $\bar{X} \subset \bar{S}$.

Since $\bar{S} \subset \hat{S}$ it follows that $\bar{S}=\hat{S}=\bar{X}$. ∎

It immediately follows from Theorem 2 that when f is quadratic, and X is polyhedral, then $\bar{X}$ is also polyhedral. Thus we have the following consequences which can also be derived by other simple direct arguments.

3. Corollary (Polyhedral characterization of the solution set of a convex quadratic program) Let X be a convex set in R^n. Consider the quadratic program

$$\min_{x \in X} f(x) := \min_{x \in X} \frac{1}{2}\, xCx + dx \tag{7}$$

where C is an $n \times n$ symmetric positive semidefinite matrix, $d \in R^n$ and let $\bar{x}$ be any point in the solution set $\bar{X}$ of $\min\limits_{x \in X} f(x)$. Then $\bar{X}$ can be characterized as follows

$$\begin{aligned} \bar{X} &= \bar{T} := \{x \mid x \in X,\ d(x-\bar{x})=0,\ C(x-\bar{x})=0\} \\ \bar{X} &= \hat{T} := \{x \mid x \in X,\ d(x-\bar{x}) \leq 0,\ C(x-\bar{x})=0\} \end{aligned} \tag{8}$$

Consequently $\bar{X}$ is a polyhedral set when X is polyhedral.

Proof Follows from Theorem 2 by noting that for x in $\bar{S}$, $\hat{S}$, $\bar{T}$ or $\hat{T}$

$$\nabla f(\bar{x})(x-\bar{x}) = (\bar{x}C+d)(x-\bar{x}) = d(x-\bar{x}). \quad ∎$$

We now give a polyhedral characterization of the solution set of the monotone linear complementarity problem [3] which is simpler than the Adler-Gale characterization [1] and which is very useful in deriving error bounds for monotone linear complementarity problems [7]. Consider the linear complementarity problem

$$Mx+q \geq 0,\ x \geq 0,\ x(Mx+q)=0 \tag{9}$$

where M is an $n \times n$ positive semidefinite matrix, not necessarily symmetric, and $q \in R^n$. We consider the equivalent formulation

$$\min \{x(Mx+q) \mid Mx+q \geq 0,\ x \geq 0\} = 0 \tag{10}$$

and apply Corollary 3 to obtain the following characterization of the solution set of (10) or equivalently of (9).

4. Corollary (Polyhedral characterization of the solution set of a monotone linear complementarity problem) Let M be positive semidefinite and let $\bar{x}$ be any point in the solution set $\bar{X}$ of the linear complementarity problem (9). Then

$$\begin{aligned} \bar{X} = \bar{U} &:= \{x | Mx + q \geq 0,\ x \geq 0,\ q(x - \bar{x}) = 0,\ (M + M^T)(x - \bar{x}) = 0\} \\ \bar{X} = \hat{U} &:= \{x | Mx + q \geq 0,\ x \geq 0,\ q(x - \bar{x}) \leq 0,\ (M + M^T)(x - \bar{x}) = 0\} \end{aligned} \tag{11}$$

We conclude with a boundedness characterization of the solution set of the monotone linear complementarity problem (9) and the convex quadratic program

$$\min \frac{1}{2} xCx + dx \quad \text{subject to } Ax \geq b,\ x \geq 0 \tag{12}$$

where C and d are as in (7), A is an $m \times n$ matrix and b is in R^m.

5. Theorem (Boundedness of the solution set of a convex quadratic program) Let C be positive semidefinite and let the quadratic program (12) have a nonempty solution set $\bar{X}$. $\bar{X}$ is bounded if and only if either

$$\begin{aligned} &\bar{Y} := \{y | Ay \geq 0,\ 0 \neq y \geq 0,\ dy = 0,\ Cy = 0\} = \phi \\ &\text{or} \\ &\hat{Y} := \{y | Ay \geq 0,\ 0 \neq y \geq 0,\ dy \leq 0,\ Cy = 0\} = \phi \end{aligned} \tag{13}$$

Proof (Sufficiency) By Corollary 3 the solution set $\bar{X}$ is given by $\bar{X} = \bar{T}$. Thus if $\bar{X}$ unbounded, there exists a sequence $\{x^i\}$, such that $x^i \neq 0$ and $\|x^i\| \to \infty$. Any accumulation point y of $\dfrac{x^i}{\|x^i\|}$ is in $\bar{Y} \subset \hat{Y}$, thus contradicting $\bar{Y} = \phi$ or $\hat{Y} = \phi$.

(Necessity) Suppose that $y \in \bar{Y} \neq \phi$ and $\bar{x} \in \bar{X}$. Then for any positive λ, $x(\lambda) := \bar{x} + \lambda y \in \bar{T} = \bar{X}$ and $\|x(\lambda)\| \to \infty$ as $\lambda \to \infty$. This however contradicts the boundedness of $\bar{X}$. ∎

Theorem 5 can be also derived by invoking a result of convex analysis [9, Theorem 8.4]. We just note that $\bar{Y} \cup \{0\}$ and $\hat{Y} \cup \{0\}$ are the recession cones of $\bar{T}$ and $\hat{T}$ respectively and by Theorem 8.4 [9], the latter are bounded if and only if the former consist of $\{0\}$ alone. It is interesting to note that the recession cone does not depend on either the solution point nor the right hand side b of the constraint $Ax \geq b$. We can thus say that the solution set of a convex quadratic program (12) is bounded (or empty) for each right hand side b if and only if it is nonempty and bounded for some b, in which case the recession cone of the solution set contains the null element only. This latter characterization which is equivalent to (13) can be verified by solving a single linear program as follows

$$\max\{ey|Ay \geq 0,\ y \geq 0,\ dy = 0,\ Cy = 0\} = 0 \tag{14}$$

where e is a vector of ones in R^n. Thus (13) and (14) are equivalent to each other.

For the monotone linear complementarity problem (9) a slightly stronger boundedness result than Theorem 5 is possible if the characterization $\bar{X} = \hat{U}$ is used. This is because of the special structure of the quadratic program (10). The boundedness characterization here implies nonemptiness of the solution set of the complementarity problem as well. We thus have the following.

6. Theorem (Nonemptiness and boundedness of the solution set of a monotone linear complementarity problem) Let M be positive semidefinite. The solution set $\bar{X}$ of the linear complementarity problem (9) is nonempty and bounded if and only if

$$\hat{Z} := \{z|Mz \geq 0,\ 0 \neq z \geq 0,\ qz \leq 0,\ (M + M^T)z = 0\} = \phi \tag{15}$$

Proof The proof of the necessity and sufficiency of (15) under the assumption that $\bar{X} \neq \phi$ is similar to the proof of Theorem 5 or because $\hat{Z} + \{0\}$ is the recession cone of $\hat{U}$. We only need to show that (15) implies that $\bar{X} \neq \phi$. The solution set $\bar{X}$ is empty if and only if [3] there is no $x \geq 0$ such that $Mx + q \geq 0$. This, by Motzkin's theorem of the alternative [5], is equivalent to the existence of a u in R^n such that

$$M^T u \leq 0,\ qu < 0,\ u \geq 0$$

But by the positive semidefiniteness of M this is equivalent to the existence of a u satisfying $(M + M^T)u = 0$, $Mu = -M^T u \geq 0$, $qu < 0$, $u \geq 0$ which contradicts (15). ∎

Theorem 6 is essentially contained in [6, Theorem 2] which was derived in a different manner for the broader class of copositive-plus matrices [3].

In the last part of the paper we extend Lemma 1 and Theorem 2 to nonsmooth convex functions. Professor A. Ben-Tal pointed out to the author that the constant gradient condition of Lemma 1 can be replaced by the nonemptiness of the intersection of the subdifferentials of each pair of points in the optimal solution set. (See (18) below.) This fact as well as our more general extension (Lemma 1a below) follow from a slightly extended lemma of Demyanov and Vasilev [4, Lemma 5.8, p. 63].

7. Lemma [4] Let f be a convex function on R^n, let $\partial f(x)$ denote its subdifferential at x, let $x^1 \neq x^2$, let $\lambda \in (0,1)$ and let $x^\lambda := (1-\lambda)x^1 + \lambda x^2$. Then

$$\partial f(x^1) \cap \partial f(x^2) \neq \phi \Longrightarrow \begin{cases} f(x^\lambda) = (1-\lambda)f(x^1) + \lambda f(x^2) \\ \partial f(x^1) \cap \partial f(x^2) \subset \partial f(x^\lambda) \end{cases} \tag{16}$$

$$\partial f(x^\lambda) = \partial f(x^1) \cap \partial f(x^2) \neq \phi \Longleftarrow f(x^\lambda) = (1-\lambda)f(x^1) + \lambda f(x^2) \tag{17}$$

The slight extension consists of $\partial f(x^1) \cap \partial f(x^2) \subset \partial f(x^\lambda)$ in (16), and $\partial f(x^\lambda) = \partial f(x^1) \cap \partial f(x^2)$ in (17), both of which follow directly from the proof of [4, pp. 63-65].

Now if we let $\bar{X}$ be the solution set of $\min_{x \in X} f(x)$ where X is a convex set in R^n and f is a convex function on R^n then

$$\partial f(x^1) \cap \partial f(x^2) \neq \phi \quad \forall x^1,\, x^2 \in \bar{X} \tag{18}$$

follows immediately from (17) because $(1-\lambda)x^1 + \lambda x^2 \in \bar{X}$ for $\lambda \in (0,1)$. In fact we state and prove an extension of Lemma 1 which subsumes (18).

1a. Lemma Let $\bar{X}$ be the nonempty solution set of $\min_{x \in X} f(x)$ where X is a convex set in R^n and f is a convex function on R^n. Let $\bar{x}$ be in the relative interior of a convex subset $\tilde{X}$ of $\bar{X}$. Then

$$\partial f(\bar{x}) = \bigcap_{x \in \tilde{X}} \partial f(x)$$

and consequently $\partial f(x)$ is constant on $ri(\bar{X})$.

Proof $\partial f(\bar{x}) \subset \bigcap_{x \in \tilde{X}} \partial f(x)$: If this inclusion fails then $\partial f(\bar{x}) \not\subset \partial f(x^1)$ for some $x^1 \in \tilde{X}$ and $x^1 \neq \bar{x}$. Since $\bar{x} \in ri(\tilde{X})$, let $\bar{x} = (1-\lambda)x^1 + \lambda x^2$ for some $x^2 \in \tilde{X}$ and $\lambda \in (0,1)$. Hence $f(\bar{x}) = (1-\lambda)f(x^1) + \lambda f(x^2)$, and by (17) of Lemma 7 we have that $\partial f(\bar{x}) = \partial f(x^1) \cap \partial f(x^2)$ contradicting $\partial f(\bar{x}) \not\subset \partial f(x^1)$.

$\bigcap_{x\in\tilde{X}} \partial f(x) \subset \partial f(\bar{x})$: If this inclusion fails then for some $v \in \bigcap_{x\in\tilde{X}} \partial f(x)$, $v \notin \partial f(\bar{x})$. Since $\bar{x} \in ri(\tilde{X})$, take $x^1, x^2 \in \tilde{X}$ such that $\bar{x} = (1-\lambda)x^1 + \lambda x^2$ and $\lambda \in (0,1)$. It follows by (16) of Lemma 7 that $v \in \partial f(x^1) \cap \partial f(x^2) \subset \partial f(\bar{x})$ contradicting $v \notin \partial f(\bar{x})$. ∎

We conclude by extending Theorem 2 to nonsmooth convex functions by means of Lemma 1a.

2a. Theorem (Characterization of the solution set of a nonsmooth convex program) Let $\bar{X}$ be the nonempty solution set of $\min_{x\in X} f(x)$, where X is a convex set in R^n and f is a convex function on R^n. Let $\bar{x} \in \bar{X}$ and let $\tilde{x} \in ri(\bar{X})$. Then

$$\bar{X} = \bar{S} = \hat{S} = \bar{S}_0 = \hat{S}_0 = \bar{S}_1 = \hat{S}_1$$

where

$$\bar{S} := \{x | x \in X,\ v(x-\bar{x}) = 0 \text{ for some } v \in \partial f(\bar{x}) \cap \partial f(x)\}$$

$$\hat{S} := \{x | x \in X,\ v(x-\bar{x}) \leq 0 \text{ for some } v \in \partial f(\bar{x}) \cap \partial f(x)\}$$

$$\bar{S}_0 := \{x | x \in X,\ v(x-\tilde{x}) = 0 \text{ for some } v \in \partial f(\tilde{x}) \subset \partial f(x)\}$$

$$\hat{S}_0 := \{x | x \in X,\ v(x-\tilde{x}) \leq 0 \text{ for some } v \in \partial f(\tilde{x}) \subset \partial f(x)\}$$

$$\bar{S}_1 := \{x | x \in X,\ v(x-\bar{x}) = 0 \text{ for some } v \in \partial f(x)\}$$

$$\hat{S}_1 := \{x | x \in X,\ v(x-\bar{x}) \leq 0 \text{ for some } v \in \partial f(x)\}$$

Proof Since $\bar{S} \subset \hat{S} \subset \hat{S}_1$, $\bar{S} \subset \bar{S}_1 \subset \hat{S}_1$ and $\bar{S}_0 \subset \hat{S}_0 \subset \hat{S}_1$, we need only show that $\hat{S}_1 \subset \bar{X}$, $\bar{X} \subset \bar{S}$ and $\bar{X} \subset \bar{S}_0$.

$(\hat{S}_1 \subset \bar{X})$ Since for $x \in \hat{S}_1$

$$f(\bar{x}) - f(x) \geq v(\bar{x} - x) \geq 0,$$

it follows that $f(\bar{x}) \geq f(x)$, and hence $x \in \bar{X}$.

$(\bar{X} \subset \bar{S})$ For $x \in \bar{X}$ and $y := (\bar{x}+x)/2$ we have by (17) of Lemma 7 that $\partial f(y) = \partial f(x) \cap \partial f(\bar{x})$, and by the minimum principle [8, Theorem 3, p. 203] since $y \in \bar{X}$

$$v(x-y) \geq 0 \text{ and } v(\bar{x}-y) \geq 0 \text{ for some } v \in \partial f(y)$$

Hence $v \in \partial f(x) \cap \partial f(\bar{x})$, $v(x-\bar{x}) = 0$ and $x \in \bar{S}$.

$(\bar{X} \subset \bar{S}_0)$ For $x \in \bar{X}$ we have $\partial f(\tilde{x}) \subset \partial f(x)$ by Lemma 1a. Since $\tilde{x} \in ri(\bar{X})$, there exists $y \in \bar{X}$ and $\lambda \in (0,1)$ such that

$$\tilde{x} = (1-\lambda)x + \lambda y$$

Therefore by the minimum principle

$$v(x-\tilde{x}) \geq 0 \text{ and } v(y-\tilde{x}) \geq 0 \text{ for some } v \in \partial f(\tilde{x})$$

This however is equivalent for $v(x-\tilde{x}) = 0$ for some $v \in \partial f(\tilde{x})$. Hence $x \in \bar{S}_0$ and $\bar{X} \subset \bar{S}_0$. ∎

References

1. I. Adler & D. Gale: “On the solutions of the positive semidefinite complementarity problem”, Report 75-12, Operations Research Center, University of California, Berkeley, August 1975.

2. A. Auslender: “Optimisation Méthodes Numériques”, Masson, Paris 1976.

3. R. W. Cottle & G. B. Dantzig: “Complementary pivot theory in mathematical programming”, Linear Algebra and Its Applications 1, 1968, 103-125.

4. V. F. Demyanov & L. V. Vasilev: “Nondifferentiable optimization”, Optimization Software, Inc., New York 1987 (Translated from Russian).

5. O. L. Mangasarian: “Nonlinear programming”, McGraw-Hill, New York 1969.

6. O. L. Mangasarian: “Characterization of bounded solutions of linear complementarity problems”, Mathematical Programming Study 19, 1982, 153-166.

7. O. L. Mangasarian & T.-S. Shiau: “Error bounds for monotone linear complementarity problems”, Mathematical Programming 36, 1986, 81-89.

8. B. T. Polyak: “Introduction to optimization”, Optimization Software, Inc., New York 1987 (Translated from Russian).

9. R. T. Rockafellar: “Convex analysis”, Princeton University Press, Princeton, New Jersey 1970.

Acknowledgement

The author is indebted to Professor A. Ben-Tal for pointing out relation (18) which prompted him to extend Lemma 1 and Theorem 2 to the nonsmooth cases of Lemma 1a and Theorem 2a.