CHARACTERIZATION OF SOLUTION SETS OF CONVEX PROGRAMS

by

J.V. Burke and M.C. Ferris

Computer Sciences Technical Report #851

June 1989

Characterization of Solution Sets of Convex Programs*

J.V. Burke† M.C. Ferris‡

June 1989

Abstract. This paper gives several characterizations of the solution set of convex programs. No differentiability of the functions involved in the problem definition is assumed. The result is a generalization of the results given in [3]. Furthermore, the subgradients attaining the minimum principle are explicitly characterized, and this characterization is shown to be independent of any solution.

Key words. Solution sets, convex programs, nonsmooth optimization

Abbreviated title. Solution Sets of Convex Programs

*This material is based on research supported by National Science Foundation Grant DMS-8803206 and Air Force Office of Scientific Research Grants ISSA-860080 and 88NM458

†Department of Mathematics, GN-50, University of Washington, Seattle, Washington 98195

‡Computer Sciences Department, University of Wisconsin, Madison, Wisconsin 53706

The purpose of this note is to extend the results given in [3] concerning the solution set of the optimization problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in S \end{array}$$

where $f\colon \mathbb{R}^n \mapsto \overline{\mathbb{R}} := \mathbb{R} \bigcup \{-\infty, +\infty\}$ is an extended real valued function which is assumed to be proper and convex and S is a convex set in $\mathbb{R}^n$. We shall assume throughout that the solution set, which we denote by $\bar{S} := \arg\min_{x \in S} f(x)$, is nonempty.

The following notation will be used. If C is a convex set, then $N(x \mid C)$ is the normal cone to C at $x \in C$ defined by

$$N(x \mid C) := \{z \mid \langle z, c - x\rangle \leq 0, \text{ for all } y \in C\ \}.$$

The convex subdifferential of f at x, $\partial f(x)$, is given by

$$\partial f(x) := \{x^* \mid f(z) \geq f(x) + \langle x^*, z - x\rangle\}.$$

The relative interior of a convex set C is denoted by $\operatorname{ri} C$ and the effective domain of an extended real valued function f is defined as

$$\operatorname{dom} f := \{x \in \mathbb{R}^n \mid \mathrm{f(x)} < +\infty\}.$$

The following characterization of the solution set of a convex program is well–known. It is often referred to as the minimum principle.

Lemma 1 *Suppose $f\colon \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ is a proper convex function and $S \neq \emptyset$ is a convex set in $\mathbb{R}^n$ and $\operatorname{ri}\operatorname{dom} f \bigcap \operatorname{ri} S \neq \emptyset$. Then $\bar{x} \in \bar{S}$ if and only if $0 \in \partial f(\bar{x}) + N(\bar{x} \mid S)$.*

A proof of this result is given in [4, Theorem 27.4]. The result can be rewritten in the following manner, which will be more convenient for this note

$$\bar{x} \in \bar{S} \iff \partial f(\bar{x}) \bigcap -N(\bar{x} \mid S) \neq \emptyset.$$

In [3], an elegant characterization of the solution set of a differentiable convex program was given, and the fact that $\nabla f(x)$ is a constant on the solution set was also established.

Both of these results were extended to the nonsmooth case, but the corresponding results are not nearly as elegant and useful since they involve some (unspecified) subgradient and the relative interior of the solution set. Since the differentiable results are useful as a theoretical tool[1, 2], we would like to characterize the subgradients which identify the solution set and remove the relative interiority assumption. In this note we prove that this can be done and show that the differentiable results have exact analogues in the nonsmooth case. First of all, we show that the subgradients which achieve the minimum principle are a constant of the problem. Using this result, we establish a generalization of the differentiable results of [3].

In the following lemma we show that $\partial f(x) \bigcap -N(x \mid S)$ is constant on the solution set of a convex program. This result has not been given before to the best knowledge of the authors. Note that if f is differentiable, the constancy of the gradient on the solution set follows immediately from part (a) of the lemma.

Lemma 2 **(a)** *$\partial f(x) \bigcap -N(x \mid S)$ is independent of $x \in \bar{S}$.*

(b) *Let $A \subseteq S$ be a convex set with $\bar{S} \bigcap A \neq \emptyset$. Then $\partial f(x) \bigcap -N(x \mid A)$ is independent of $x \in \bar{S} \bigcap A$.*

Proof

(a) Let $\bar{x} \in \bar{S}$ and take $v \in \partial f(\bar{x}) \bigcap -N(\bar{x} \mid S)$. $-v \in N(\bar{x} \mid S)$ gives

$$\langle v, x - \bar{x} \rangle \geq 0, \text{ for all } x \in S \tag{1}$$

$v \in \partial f(\bar{x})$ implies that

$$\langle v, x - \bar{x} \rangle \leq f(x) - f(\bar{x}), \text{ for all } x. \tag{2}$$

Let $\tilde{x} \in \bar{S}$. It follows from (1) that $\langle v, \tilde{x} - \bar{x} \rangle \geq 0$ and from (2) that $\langle v, \tilde{x} - \bar{x} \rangle \leq 0$, so that

$$\langle v, \tilde{x} - \bar{x} \rangle = 0. \tag{3}$$

Substituting (3) in (1) we find $\langle v, x - \tilde{x} \rangle \geq 0$, for all $x \in S$, so $-v \in N(\tilde{x} \mid S)$. Using (3) in (2) and $\tilde{x} \in \bar{S}$ gives $\langle v, x - \tilde{x} \rangle \leq f(x) - f(\tilde{x})$, for all x, implying $v \in \partial f(\tilde{x})$. Hence $v \in \partial f(\tilde{x}) \bigcap -N(\tilde{x} \mid S)$. The result now follows since $\tilde{x}$ and $\bar{x}$ are arbitrary in $\bar{S}$.

(b) Consider

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in A \end{array}$$

. Applying part (a) to this problem gives the required result. $\square$

We remark that although the sets described above are constant on the solution set in general they are not the same set. The following example exhibits this fact.

Example 3 *Let* $f(x,y) = |x|$ *and let* $S = \{(x,y) \mid x \geq 0\}$. *Then* $\bar{S} = \{(0,y) \mid y \in \mathbb{R}\}$ *and so for* $z = (0,y)$

$$\partial f(z) \bigcap -N(z \mid S) = [0,1] \times \{0\}$$

and

$$\partial f(z) \bigcap -N(z \mid \bar{S}) = [-1,1] \times \{0\}$$

which are clearly different.

The following lemma enables us to relate the above results to those given in [3].

Lemma 4 *If* $\bar{x} \in \operatorname{ri} \bar{S}$ *then* $\partial f(\bar{x}) \subseteq -N(\bar{x} \mid \bar{S})$.

Proof Let $x^* \in \partial f(\bar{x})$, so that

$$0 \geq \langle x^*, z - \bar{x} \rangle, \quad \forall z \in \bar{S}.$$

It follows from [4, Theorem 6.4] and $\bar{x} \in \operatorname{ri} \bar{S}$ that for each $y \in \bar{S}$ there is some $\epsilon > 0$ with $\bar{x} - \epsilon(y - \bar{x}) \in \bar{S}$. Hence

$$0 \geq \langle x^*, \bar{x} - \epsilon(y - \bar{x}) - \bar{x} \rangle$$

which implies that $0 \geq \langle -x^*, y - \bar{x} \rangle$ as required. $\square$

In [3, Lemma 1a], Mangasarian showed that the subdifferential is constant on the relative interior of the solution set of a convex program. This follows from the above result, since

Lemma 4 shows that $\partial f(z) = \partial f(z) \bigcap -N(z \mid \bar{S})$ on the relative interior of the solution set and Lemma 2(a) shows the latter set to be a constant of the problem. However, this set is not the set where the minimum principle is achieved as the example shows, that is, there are some subgradients in this set which do not achieve the minimum principle, and this is precisely the reason that the subgradient is not specified explicitly in [3, Theorem 1a].

In the following theorem we give another characterization of the solution set. In contrast to [3, Theorem 1a] we describe precisely the subgradients which are used to form our characterization.

Theorem 5 *Suppose* $\bar{x} \in \bar{S}$. *Let*

$$\tilde{S} := \left\{ x \in S \,\middle|\, \partial f(x) \bigcap -N(x \mid S) = \partial f(\bar{x}) \bigcap -N(\bar{x} \mid S) \right\}$$

and let A *be a convex set with* $\bar{S} \subseteq A \subseteq S$ *and*

$$\hat{S}_A := \left\{ x \in A \,\middle|\, \partial f(x) \bigcap -N(x \mid A) = \partial f(\bar{x}) \bigcap -N(\bar{x} \mid A) \right\}.$$

Then $\bar{S} = \tilde{S} = \hat{S}_A$.

Proof

$\bar{S} \subset \tilde{S}$: Let $z \in \bar{S}$. Then $z \in S$ and by Lemma 2(a) it follows that $z \in \tilde{S}$.

$\tilde{S} \subset \hat{S}_A$: Let $z \in \tilde{S}$. Since $\bar{x} \in \bar{S}$, $\partial f(\bar{x}) \bigcap -N(\bar{x} \mid S) \neq \emptyset$, so $\exists v \in \partial f(z) \bigcap -N(z \mid S)$. Therefore $z \in \bar{S}$ since $0 \in \partial f(z) + N(z \mid S)$ and so from Lemma 2(b)

$$\partial f(z) \bigcap -N(z \mid A) = \partial f(\bar{x}) \bigcap -N(\bar{x} \mid A)$$

which implies that $z \in \hat{S}_A$.

$\hat{S}_A \subset \bar{S}$: Let $z \in \hat{S}_A$. Then $z \in S$ and note that $\partial f(\bar{x}) \bigcap -N(\bar{x} \mid S) \neq \emptyset$, since $\bar{x} \in \bar{S}$ which implies that $\partial f(\bar{x}) \bigcap -N(\bar{x} \mid A) \neq \emptyset$, since $A \subseteq S$. Hence $\exists v \in \partial f(z) \bigcap -N(z \mid A)$. But $v \in \partial f(z)$ implies $f(y) \geq f(z) + \langle v, y - z \rangle$, for all y and $-v \in N(z \mid A)$ implies

$\langle -v, y - z\rangle \leq 0$, for all $y \in A$. Therefore

$$\begin{aligned} f(y) &\geq f(z) + \langle v, y - z\rangle && \text{for all } y \\ &\geq f(z) && \text{for all } y \in A\ . \end{aligned}$$

Thus $z \in \bar{S}$ since z is feasible and has objective value at least as good as the optimal value, since $\bar{S} \subseteq A$. $\square$

The differentiable result now follows immediately, since by Theorem 5, $\bar{S} = \hat{S}_A$ with $A = \bar{S}$, and so if $x \in \bar{S}$ then $\partial f(x) \subseteq \{\nabla f(\bar{x})\}$ and $-\nabla f(\bar{x}) \in N(x \mid \bar{S})$.

Corollary 6 ([3, Theorem 1]) *Let f be a differentiable convex function and $\bar{x} \in \bar{S}$. Then*

$$\bar{S} = \{x \in S \mid \langle \nabla f(\bar{x}), x - \bar{x}\rangle \leq 0,\ \ \nabla f(x) = \nabla f(\bar{x})\}\,.$$

Using Theorem 5, this corollary has the following generalization.

Corollary 7 *Let $A \subset S$ be such that $A \bigcap \bar{S} \neq \emptyset$. Choose $\bar{x} \in \bar{S} \bigcap A$ and set*

$$\hat{S}_A := \left\{ x \in A \,\middle|\, \partial f(x) \bigcap -N(x \mid A) = \partial f(\bar{x}) \bigcap -N(\bar{x} \mid A) \right\}.$$

Then $\bar{S} \bigcap A = \hat{S}_A = \arg\min_{x \in A} f(x)$.

References

[1] J.V. Burke and M.C. Ferris. The sharpness of functions on sets. In preparation, 1989.

[2] M.C. Ferris and O.L. Mangasarian. Finite perturbation of convex programs. Technical Report 802, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin 53706, October 1988. Revised, June 1989.

[3] O.L. Mangasarian. A simple characterization of solution sets of convex programs. *Operations Research Letters*, 7(1):21–26, 1988.

[4] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, NJ, 1970.