

FINITE PERTURBATION OF CONVEX PROGRAMS

by

M. C. Ferris and O. L. Mangasarian

Computer Sciences Technical Report #802

October 1988

Finite Perturbation of Convex Programs

M.C. Ferris and O.L. Mangasarian*

Computer Sciences Department

University of Wisconsin

Madison, Wisconsin 53706

October 1988

Abstract. This paper concerns a characterization of the finite perturbation property of a convex program. When this property holds, finite perturbation of the objective function of a convex program leads to a solution of the original problem which minimizes the perturbation function over the set of solutions of the convex program. This generalizes an important property of least norm solutions of linear programs which plays a key role in powerful algorithms for solving large sparse linear programs. It also generalizes a finite termination property of the proximal point algorithm.

Key words. Exact penalty, finite perturbation, convex programming

Abbreviated title. Finite Perturbation

*This material is based on research supported by National Science Foundation Grants DCR-8521228 and CCR-8723091 and Air Force Office of Scientific Research Grant AFOSR-86-0172

1 Introduction

The motivation of this work comes from the desire to extend the finite perturbation property of linear programs[7] to convex programs. This property, which is naturally possessed by linear programs, plays an important role in the iterative solution of very large-scale linear programs[4, 2] and in the finite termination of the proximal point algorithm for linear programs[8]. Basically, this property consists of the equivalence of minimizing a convex function on the **solution set** of a linear program to the problem of minimizing a perturbation of the objective function of the linear program by **any** sufficiently small positive multiple of the same convex function.

The paper is organised as follows. In Section 2 we derive an equivalence for the finite perturbation property under assumed differentiability of the functions concerned. Section 3 generalizes these conditions to a nondifferentiable setting.

We shall be concerned principally with the convex programming problem

$$\underset{x \in S}{\text{minimize}} \quad f(x) \tag{1}$$

where f is a lower semicontinuous, convex function defined on \mathbb{R}^n , having values in \mathbb{R} and S is a closed, convex set in \mathbb{R}^n . We write \bar{S} for the optimal solution set of (1), $\bar{S} := \arg \min_{x \in S} f(x)$ and assume this set to be nonempty.

The notation we use is for the most part standard. The following partial list is provided for the reader's convenience. Superscripts are used to distinguish between vectors, e.g., x^1 , x^2 , etc. and \langle, \rangle is used to denote the inner product. If K is a cone, then we define the polar of K , by $K^\circ := \{x^* \in \mathbb{R}^n \mid \langle x^*, x \rangle \leq 0, \forall x \in K\}$. If C is a convex set, then $\text{ri } C$ is the relative interior of the set C , $N(x \mid C) := \{x^* \in \mathbb{R}^n \mid \langle x^*, c - x \rangle \leq 0, \forall c \in C\}$ is the normal cone to C at $x \in C$ and $T(x \mid C) := N(x \mid C)^\circ$ is the tangent cone to C at $x \in C$. For $f: \mathbb{R}^n \mapsto \mathbb{R}$, f convex, $\partial f(x)$ is the subdifferential of f at x ,

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

The symbol $\|\cdot\|$ denotes the 2-norm. For a real number λ , λ_+ denotes $\max\{\lambda, 0\}$.

2 Finite perturbation with differentiability

We begin with a simple result which relates a solution of a finitely perturbed nonlinear program to a solution of the original problem which minimizes the perturbation function. Note that no convexity is needed here.

Lemma 1 *Let $S \subseteq \mathbb{R}^n$, $f: \mathbb{R}^n \mapsto \mathbb{R}$ and assume $\emptyset \neq \bar{S} := \operatorname{argmin}_{x \in S} f(x)$. Let $\theta: \mathbb{R}^n \mapsto \mathbb{R}$ and suppose $\emptyset \neq \operatorname{argmin}_{x \in \bar{S}} \theta(x)$. If*

$$\bar{x} \in \operatorname{argmin}_{x \in S} f(x) + \epsilon \theta(x) \quad \forall \epsilon \in (0, \bar{\epsilon}]$$

then $\bar{x} \in \operatorname{argmin}_{x \in \bar{S}} \theta(x)$.

Proof Consider the following well-defined problem which is the equivalent of $\min_{x \in \bar{S}} \theta(x)$

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & f(x) \leq \bar{f}, \quad x \in S \end{array}$$

where $\bar{f} := \min_{x \in S} f(x)$, and consider the associated exterior penalty function problem

$$\min \left\{ \epsilon \theta(x) + (f(x) - \bar{f})_+ \mid x \in S \right\} = \min \left\{ \epsilon \theta(x) + (f(x) - \bar{f}) \mid x \in S \right\}$$

By assumption \bar{x} solves this penalty problem for all $\epsilon \in (0, \bar{\epsilon}]$. Hence, by [5, Theorem 2.5]

$$\bar{x} \in S, \quad f(\bar{x}) \leq \bar{f} \text{ and } \theta(\bar{x}) \leq \min \left\{ \theta(x) \mid x \in S, \quad f(x) \leq \bar{f} \right\}$$

Hence $\bar{x} \in \operatorname{argmin}_{x \in \bar{S}} \theta(x)$. ■

We now show that a solution of a finitely perturbed differentiable convex program is also a solution of the unperturbed problem with a linearized objective function around any of its solution points and which minimizes the perturbation function.

Theorem 2 Let $f: \mathbb{R}^n \mapsto \mathbb{R}$, $\theta: \mathbb{R}^n \mapsto \mathbb{R}$, $g: \mathbb{R}^n \mapsto \mathbb{R}^m$ be convex and differentiable on \mathbb{R}^n and let $S := \{x \mid g(x) \leq 0\}$ and $\emptyset \neq \bar{S} := \operatorname{argmin}_{x \in S} f(x)$. If

$$\bar{x} \in \operatorname{argmin}_{x \in S} f(x) + \epsilon \theta(x) \quad \text{for some } \epsilon > 0 \quad (2)$$

and $\emptyset \neq \operatorname{int} S := \{x \mid g(x) < 0\}$ then

$$\bar{x} \in \operatorname{argmin} \{\theta(x) \mid \langle \nabla f(\hat{x}), x - \hat{x} \rangle \leq 0, \ g(x) \leq 0\}$$

for any $\hat{x} \in \bar{S}$.

Proof By the Karush–Kuhn–Tucker conditions for (2), there exists $u(\epsilon) \in \mathbb{R}^m$ such that

$$\left. \begin{aligned} \nabla f(\bar{x}) + \epsilon \nabla \theta(\bar{x}) + u(\epsilon)^T \nabla g(\bar{x}) &= 0 \\ \langle u(\epsilon), g(\bar{x}) \rangle &= 0 \\ g(\bar{x}) &\leq 0 \\ u(\epsilon) &\geq 0 \end{aligned} \right\} \quad \forall \epsilon \in (0, \bar{\epsilon}] \quad (3)$$

By Lemma 1, $\bar{x} \in \operatorname{argmin}_{x \in \bar{S}} \theta(x)$. Hence $\bar{x} \in \bar{S}$ and by [6, Theorem 1], $\nabla f(\bar{x}) = \nabla f(\hat{x})$ and $\langle \nabla f(\bar{x}), \bar{x} \rangle = \langle \nabla f(\hat{x}), \hat{x} \rangle$. But (3) are also the Karush–Kuhn–Tucker conditions for

$$\min \{\theta(x) \mid \langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq 0, \ g(x) \leq 0\} = \min \{\theta(x) \mid \langle \nabla f(\hat{x}), x - \hat{x} \rangle \leq 0, \ g(x) \leq 0\}$$

Since θ and g are convex, it follows that

$$\bar{x} \in \operatorname{argmin} \{\theta(x) \mid \langle \nabla f(\hat{x}), x - \hat{x} \rangle \leq 0, \ g(x) \leq 0\}$$

which is the result required. ■

We establish a sort of converse to Theorem 2 by showing that if \bar{x} minimizes the perturbation function over both the solution set of the unperturbed convex program and over a linearization of the convex program, then \bar{x} is also a solution of the finitely perturbed convex program.

Theorem 3 *Let $f: \mathbb{R}^n \mapsto \mathbb{R}$, $g: \mathbb{R}^n \mapsto \mathbb{R}^m$ be convex and differentiable on \mathbb{R}^n , let $\theta: \mathbb{R}^n \mapsto \mathbb{R}$ be strongly convex and differentiable on \mathbb{R}^n and let*

$$S := \{x \mid g(x) \leq 0\}, \quad \bar{S} := \operatorname{argmin}_{x \in S} f(x) \text{ and } \bar{x} = \operatorname{argmin}_{x \in \bar{S}} \theta(x)$$

Define

$$LS(\bar{x}) := \{x \mid g(\bar{x}) + \nabla g(\bar{x})(x - \bar{x}) \leq 0\}$$

Assume

$$\bar{x} = \operatorname{argmin} \{\theta(x) \mid \langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq 0, \quad x \in LS(\bar{x})\} \quad (4)$$

and that S satisfies a constraint qualification. Then

$$\bar{x} = \operatorname{argmin}_{x \in S} f(x) + \epsilon \theta(x) \quad \forall \epsilon \in (0, \bar{\epsilon}]$$

Proof It follows from $\bar{x} \in \bar{S}$ and the constraint qualification that

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \forall x \in LS(\bar{x})$$

and hence that

$$\{x \mid \langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq 0, \quad x \in LS(\bar{x})\} = \operatorname{argmin}_{x \in LS(\bar{x})} \langle \nabla f(\bar{x}), x - \bar{x} \rangle$$

Invoking [7, Theorem 4] we see that for all $\epsilon \in (0, \bar{\epsilon}]$, for some $\bar{\epsilon} > 0$

$$\operatorname{argmin} \left\{ \theta(x) \mid \begin{array}{l} \langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq 0 \\ x \in LS(\bar{x}) \end{array} \right\} = \operatorname{argmin}_{x \in LS(\bar{x})} \langle \nabla f(\bar{x}), x - \bar{x} \rangle + \epsilon \theta(x)$$

It follows by (4) that \bar{x} is the unique solution of the last problem above, for $\epsilon \in (0, \bar{\epsilon}]$. Since $\bar{x} \in S \subseteq LS(\bar{x})$ we have

$$\bar{x} = \operatorname{argmin}_{x \in S} \langle \nabla f(\bar{x}), x - \bar{x} \rangle + \epsilon \theta(x) \quad \forall \epsilon \in (0, \bar{\epsilon}]$$

Since S satisfies a constraint qualification, \bar{x} and some $u(\epsilon)$ satisfy the Karush–Kuhn–Tucker conditions for this problem, namely

$$\left. \begin{aligned} \nabla f(\bar{x}) + \epsilon \nabla \theta(\bar{x}) + u(\epsilon)^T \nabla g(\bar{x}) &= 0 \\ \langle u(\epsilon), g(\bar{x}) \rangle &= 0 \\ g(\bar{x}) &\leq 0 \\ u(\epsilon) &\geq 0 \end{aligned} \right\} \quad \forall \epsilon \in (0, \bar{\epsilon}]$$

But these are precisely the Karush–Kuhn–Tucker conditions for

$$\underset{x \in S}{\text{minimize}} f(x) + \epsilon \theta(x) \quad \forall \epsilon \in (0, \bar{\epsilon}]$$

Hence $\bar{x} \in \underset{x \in S}{\text{argmin}} f(x) + \epsilon \theta(x)$, for $\epsilon \in (0, \bar{\epsilon}]$. ■

For the case when the feasible region is polyhedral, it is possible to combine the necessary conditions for the finite perturbation property (Lemma 1 and Theorem 2) and the sufficient condition (Theorem 3) into the following necessary and sufficient condition.

Theorem 4 *Let the definitions of Theorem 3 hold with g being linear on \mathbb{R}^n . Then*

$$\bar{x} = \underset{x \in S}{\text{argmin}} f(x) + \epsilon \theta(x) \quad \forall \epsilon \in (0, \bar{\epsilon}]$$

if and only if

$$\bar{x} = \underset{x \in \bar{S}}{\text{argmin}} \theta(x) = \underset{x \in \bar{S}}{\text{argmin}} \{ \theta(x) \mid \langle \nabla f(\hat{x}), x - \hat{x} \rangle \leq 0, \ x \in S \}$$

where \hat{x} is any point in \bar{S} .

We give the following application of Theorem 4 to the one-step termination of the proximal point algorithm.

Theorem 5 *Let $f: \mathbb{R}^n \mapsto \mathbb{R}$ be convex and differentiable on \mathbb{R}^n , let $g: \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear on \mathbb{R}^n , let $x^0 \in \mathbb{R}^n$ and let*

$$S := \{x \mid g(x) \leq 0\}, \quad \emptyset \neq \bar{S} := \underset{x \in S}{\text{argmin}} f(x)$$

Then

$$\bar{x} = \operatorname{argmin}_{x \in S} f(x) + \frac{\epsilon}{2} \|x - x^0\|^2 \quad \forall \epsilon \in (0, \bar{\epsilon}]$$

if and only if

$$\bar{x} = \operatorname{argmin}_{x \in \bar{S}} \|x - x^0\|^2 = \operatorname{argmin} \left\{ \|x - x^0\|^2 \mid \langle \nabla f(\hat{x}), x - \hat{x} \rangle \leq 0, \ x \in S \right\}$$

where \hat{x} is any point in \bar{S} .

3 Finite perturbation without differentiability

In this section we generalize the results of the previous section to the nondifferentiable case. We first give some elementary lemmas which will be needed in the proofs of the following results.

Lemma 6 *If S is convex and $x \in S$ then*

$$N(x \mid S) = N(x \mid x + T(x \mid S))$$

Proof

$$\begin{aligned} x^* \in N(x \mid S) &\iff x^* \in N(x \mid S)^{\circ\circ} \iff x^* \in T(x \mid S)^{\circ} \\ &\iff \langle x^*, u \rangle \leq 0, \text{ for all } u \in T(x \mid S) \\ &\iff \langle x^*, z - x \rangle \leq 0, \text{ for all } z \in x + T(x \mid S) \\ &\iff x^* \in N(x \mid x + T(x \mid S)) \end{aligned}$$

■

Lemma 7 *Let A, B be convex sets and let $x \in A \cap B$. Then*

$$N(x \mid A) + N(x \mid B) \subseteq N(x \mid A \cap B)$$

Proof Let $u \in N(x \mid A)$, $v \in N(x \mid B)$, then $0 \geq \langle u, z - x \rangle$, for all $z \in A$ and $0 \geq \langle v, z - x \rangle$, for all $z \in B$ so that

$$0 \geq \langle u + v, z - x \rangle \quad \text{for all } z \in A \cap B$$

giving $u + v \in N(x \mid A \cap B)$ as required. ■

The following result extends Theorem 2 to the nondifferentiable case.

Theorem 8 *Let $f: \mathbb{R}^n \mapsto \mathbb{R}$ and $\theta: \mathbb{R}^n \mapsto \mathbb{R}$ be convex functions. Suppose $\emptyset \neq \bar{S} := \operatorname{argmin}_{x \in S} f(x)$. If*

$$\bar{x} \in \operatorname{argmin}_{x \in S} f(x) + \epsilon \theta(x) \quad \text{for some } \epsilon > 0 \tag{5}$$

then

$$\bar{x} \in \operatorname{argmin} \{ \theta(x) \mid \langle v^*, x - \bar{x} \rangle \leq 0, \ x \in \bar{x} + T(\bar{x} \mid S) \}$$

for some $v^ \in \partial f(\bar{x})$.*

Proof

$$\begin{aligned} (5) &\implies 0 \in \partial f(\bar{x}) + \epsilon \partial \theta(\bar{x}) + N(\bar{x} \mid S) \\ &\implies 0 \in v^* + \epsilon \partial \theta(\bar{x}) + N(\bar{x} \mid S) && (\text{for some } v^* \in \partial f(\bar{x})) \\ &\implies 0 \in \frac{v^*}{\epsilon} + \partial \theta(\bar{x}) + N(\bar{x} \mid S) \\ &\implies 0 \in \partial \theta(\bar{x}) + \frac{v^*}{\epsilon} + N(\bar{x} \mid \bar{x} + T(\bar{x} \mid S)) && (\text{by Lemma 6}) \\ &\implies 0 \in \partial \theta(\bar{x}) + N(\bar{x} \mid \{x \mid \langle v^*, x - \bar{x} \rangle \leq 0\}) + N(\bar{x} \mid \bar{x} + T(\bar{x} \mid S)) \\ &\hspace{15em} (\text{since } \frac{v^*}{\epsilon} \in N(\bar{x} \mid \{x \mid \langle v^*, x - \bar{x} \rangle \leq 0\})) \\ &\implies \bar{x} \in \operatorname{argmin} \{ \theta(x) \mid \langle v^*, x - \bar{x} \rangle \leq 0, \ x \in \bar{x} + T(\bar{x} \mid S) \} && (\text{by Lemma 7}) \end{aligned}$$

■

The following lemmas are needed in the proof of Theorem 11 and their proofs are found in the literature.

Lemma 9 ([1, Theorem 2.3]) *Let A, B be convex sets with $A \subseteq B$. Let $x \in \operatorname{ri} A$ and $y \in A$. Then $N(x \mid B) \subseteq N(y \mid B)$.*

Lemma 10 ([6, Lemma 1a]) *Let $\bar{S} = \operatorname{argmin}_{x \in S} f(x)$, where S is a convex subset of \mathbb{R}^n and $f: \mathbb{R}^n \mapsto \mathbb{R}$ is convex. Let $x \in \operatorname{ri} \bar{S}$. Then $\partial f(x) \subseteq \bigcap_{y \in \bar{S}} \partial f(y)$.*

Using Lemma 10, we can define the **effective subdifferential** of f at $x \in \bar{S}$ as

$$\operatorname{eff} \partial f(x) := \bigcap_{y \in \bar{S}} \partial f(y) \neq \emptyset \quad (6)$$

The nonemptiness of $\operatorname{eff} \partial f(x)$ follows from Lemma 10 and the nonemptiness of $\operatorname{ri} \bar{S}$. We now extend Theorem 3 to the nondifferentiable case. In the theorem below, we assume that $T(\bar{x} \mid S)$ is polyhedral. For the set given in the previous section, this is essentially a constraint qualification, see for example the Guignard constraint qualification[3].

Theorem 11 *Let $f: \mathbb{R}^n \mapsto \mathbb{R}$ be convex, let $\theta: \mathbb{R}^n \mapsto \mathbb{R}$ be strongly convex and let S be a convex set such that*

$$\bar{S} := \operatorname{argmin}_{x \in S} f(x) \text{ and } \bar{x} = \operatorname{argmin}_{x \in \bar{S}} \theta(x)$$

Assume that $T(\bar{x} \mid S)$ is polyhedral and

$$\bar{x} = \operatorname{argmin} \{ \theta(x) \mid \langle v^*, x - \bar{x} \rangle \leq 0, \ x \in \bar{x} + T(\bar{x} \mid S) \} \quad (7)$$

where

$$v^* \in \operatorname{eff} \partial f(\bar{x}) \cap -N(\bar{x} \mid S) \quad (8)$$

Then

$$\bar{x} = \operatorname{argmin}_{x \in S} f(x) + \epsilon \theta(x) \quad \forall \epsilon \in (0, \bar{\epsilon}]$$

Proof We first prove the existence of a v^* satisfying (8). Since $\bar{S} \neq \emptyset$ it follows that $\operatorname{ri} \bar{S} \neq \emptyset$. Therefore, take $\tilde{x} \in \operatorname{ri} \bar{S}$ so that by Lemma 10

$$\operatorname{eff} \partial f(x) = \partial f(\tilde{x})$$

Since $\tilde{x} \in \bar{S}$ there is a $v^* \in \partial f(\tilde{x})$ and $x^* \in N(\tilde{x} \mid S)$ satisfying $0 = v^* + x^*$. By Lemma 9, $x^* \in N(\tilde{x} \mid S) \subseteq N(\bar{x} \mid S)$ so that v^* satisfies (8). We take any v^* satisfying (8).

From the definition, $-v^* \in N(\bar{x} \mid S)$ and hence by Lemma 6, $-v^* \in N(\bar{x} \mid \bar{x} + T(\bar{x} \mid S))$. It follows that

$$\langle v^*, x - \bar{x} \rangle \geq 0 \quad \text{for all } x \in \bar{x} + T(\bar{x} \mid S)$$

so that

$$\{x \mid \langle v^*, x - \bar{x} \rangle \leq 0, \ x \in \bar{x} + T(\bar{x} \mid S)\} = \operatorname{argmin} \{\langle v^*, x - \bar{x} \rangle \mid x \in \bar{x} + T(\bar{x} \mid S)\}$$

Invoking [7, Theorem 4] we see that

$$\operatorname{argmin} \left\{ \theta(x) \left| \begin{array}{l} \langle v^*, x - \bar{x} \rangle \leq 0 \\ x \in \bar{x} + T(\bar{x} \mid S) \end{array} \right. \right\} = \operatorname{argmin} \{\langle v^*, x - \bar{x} \rangle + \epsilon \theta(x) \mid x \in \bar{x} + T(\bar{x} \mid S)\},$$

and by (7), \bar{x} is the unique solution of the last problem above for $\epsilon \in (0, \bar{\epsilon}]$. Since $\bar{x} \in S \subseteq \bar{x} + T(\bar{x} \mid S)$ it follows that

$$\bar{x} = \operatorname{argmin}_{x \in S} \langle v^*, x - \bar{x} \rangle + \epsilon \theta(x) \quad \forall \epsilon \in (0, \bar{\epsilon}]$$

Therefore, for all $\epsilon \in (0, \bar{\epsilon}]$

$$\begin{aligned} 0 &\in v^* + \epsilon \partial \theta(\bar{x}) + N(\bar{x} \mid S) \\ &\subseteq \partial f(\bar{x}) + \epsilon \partial \theta(\bar{x}) + N(\bar{x} \mid S) \\ &= \partial(f + \epsilon \theta)(\bar{x}) + N(\bar{x} \mid S) \end{aligned}$$

Therefore $\bar{x} \in \operatorname{argmin}_{x \in S} f(x) + \epsilon \theta(x)$, for $\epsilon \in (0, \bar{\epsilon}]$. ■

Acknowledgement

The authors would like to thank Mr. D. Ralph for useful discussions during the preparation of this paper.

References

- [1] J.V. Burke and J.J. Moré. On the identification of active constraints. *SIAM Journal of Numerical Analysis*, 1988. To appear.
- [2] R. De Leone and O.L. Mangasarian. Serial and parallel solution of large scale linear programs by augmented Lagrangian successive overrelaxation. In A. Kurzhanski, K. Neumann, and D. Pallaschke, editors, *Optimization, Parallel Processing and Applications*, pages 103–124, Springer-Verlag, Berlin, 1988. Lecture Notes in Economics and Mathematical Systems 304.
- [3] M. Guignard. Generalized Kuhn–Tucker conditions for mathematical programming in a Banach space. *SIAM Journal on Control and Optimization*, 7:232–241, 1969.
- [4] O.L. Mangasarian. Normal solutions of linear programs. *Mathematical Programming Study*, 22:206–216, 1984.
- [5] O.L. Mangasarian. Some applications of penalty functions in mathematical programming. In R. Conti, E. De Giorgi, and F. Giannessi, editors, *Optimization and Related Fields*, pages 307–329, Springer-Verlag, Heidelberg, 1986. Lecture Notes in Mathematics 1190.
- [6] O.L. Mangasarian. A simple characterization of solution sets of convex programs. *Operations Research Letters*, 7(1):21–26, 1988.
- [7] O.L. Mangasarian and R.R. Meyer. Nonlinear perturbation of linear programs. *SIAM Journal on Control and Optimization*, 17(6):745–752, November 1979.
- [8] B.T. Polyak and N.V. Tretiyakov. Concerning an iterative method for linear programming and its economic interpretation. *Economics and Mathematical Methods*, 8(5):740–751, 1972. (Russian).