

NONLINEAR PERTURBATION OF LINEAR PROGRAMS*

O. L. MANGASARIAN† AND R. R. MEYER†

Abstract. The objective function of any solvable linear program can be perturbed by a differentiable, convex or Lipschitz continuous function in such a way that (a) a solution of the original linear program is also a Karush–Kuhn–Tucker point, local or global solution of the perturbed program, or (b) each global solution of the perturbed problem is also a solution of the linear program.

We are concerned here with the linear program

$$(1) \quad \text{Minimize } px \quad \text{subject to} \quad Ax \cong b,$$

where p and b are given vectors in R^n and R^m respectively and A is a given $m \times n$ real matrix. We shall assume throughout this work that this problem has a nonempty optimal solution set $\bar{S} \subset S = \{x | Ax \cong b\}$. We shall be interested in the perturbed problem $P(\varepsilon)$ defined as follows:

$$(2) \quad \text{Minimize } px + \varepsilon f(x) \quad \text{subject to} \quad Ax \cong b,$$

where $f: R^n \rightarrow R$ and ε is a nonnegative real number. For convenience we define the optimal solution set of (2) as $\bar{S}(\varepsilon)$. Note that $\bar{S}(0) = \bar{S}$. Perturbed problems such as (2) are considered in [3], [4]. In [3] it was shown that if (1) has a *unique* solution \bar{x} and f is a differentiable function at \bar{x} , then there exists a positive $\bar{\varepsilon}$ such that for all ε in $[0, \bar{\varepsilon}]$, \bar{x} satisfies the Karush–Kuhn–Tucker conditions [1], [2] for the perturbed problem (2). By considering a specific perturbation $f(x) = \frac{1}{2}x^T x$ in [4] an iterative technique is proposed for solving linear programming problems. In this work we show that, under suitable conditions, given by f there exists a positive number $\bar{\varepsilon}$ such that some solution of the linear program is a Karush–Kuhn–Tucker point or a local or global solution of the perturbed problem (2) for ε in the interval $[0, \bar{\varepsilon}]$. In Theorem 1 we show that if f is differentiable and has a bounded level set on \bar{S} then there exists a Karush–Kuhn–Tucker point of the perturbed problem (2) which also solves the original linear program (1). In Theorem 2 we indicate how the same type of perturbation applies to a nonlinear programming problem. The rest of the paper is again devoted to the perturbed linear program. In Theorem 3 we show that if f satisfies a local Lipschitz or local convexity property then there exists a solution of the linear program (1) which is a local solution of the perturbed problem (2). Among other things Theorem 4 globalizes the result of Theorem 3 and shows that for sufficiently small $\varepsilon \cong 0$ the set of optimal solutions of the perturbed problem is actually a subset of the solutions of (1). Corollary 1 deals with the case when the linear program (1) has a unique solution, while Corollary 2 treats the case when the perturbation function f is strictly convex on R^n . We begin with the first result.

THEOREM 1. *Let f be a function from R^n into R which is differentiable on the nonempty solution set \bar{S} of (1). Let either the level set $L = \{x | x \in \bar{S}, f(x) \leq \beta\}$ be nonempty and bounded for some real number β , or let θ be the minimum value of (1) and let the nonlinear program*

$$(3) \quad \text{Minimize } f(x) \quad \text{subject to} \quad Ax \cong b, \quad px \leq \bar{\theta}$$

have a Karush–Kuhn–Tucker point. Then there exists an \bar{x} in R^n and an $\bar{\varepsilon} > 0$ such that

* Received by the editors November 8, 1978 and in revised form March 25, 1979.

† Computer Sciences and Industrial Engineering Departments, University of Wisconsin—Madison, Madison, Wisconsin 53706. This research was supported in part by National Science Foundation Grant MCS74-20584 A02.

for each ϵ in $[0, \bar{\epsilon}]$ there exists a $\bar{u}(\epsilon)$ in R^m such that $(\bar{x}, \bar{u}(\epsilon))$ is a Karush-Kuhn-Tucker point of the perturbed problem (2), and \bar{x} is also a solution of the linear program (1). If in addition f is convex or pseudoconvex at \bar{x} , then \bar{x} solves the perturbed problem (2) for ϵ in $[0, \bar{\epsilon}]$.

Proof. By explicit assumption or by the boundedness of L , problem (3) has a Karush-Kuhn-Tucker point $(\bar{x}, \bar{v}, \bar{\gamma})$ in R^{n+m+1} which satisfies

$$\begin{aligned}
 \nabla f(\bar{x}) - A^T \bar{v} + \bar{\gamma} p &= 0, \\
 A\bar{x} &\geq b, \\
 p\bar{x} &= \bar{\theta}, \\
 \bar{v}(A\bar{x} - b) &= 0, \\
 \bar{v}, \bar{\gamma} &\geq 0.
 \end{aligned}
 \tag{4}$$

Since \bar{x} is also a solution of the linear program (1), there exists a \bar{w} in R^m such that

$$-A^T \bar{w} + p = 0, \quad A\bar{x} \geq b, \quad \bar{w}(A\bar{x} - b) = 0, \quad \bar{w} \geq 0.
 \tag{5}$$

Case 1: $\bar{\gamma} = 0$. From (4) and (5) we have that for any $\epsilon \geq 0$

$$\begin{aligned}
 \epsilon \nabla f(\bar{x}) - A^T(\bar{w} + \epsilon \bar{v}) + p &= 0, \\
 A\bar{x} &\geq b, \\
 (\bar{w} + \epsilon \bar{v})(A\bar{x} - b) &= 0, \\
 \bar{w} + \epsilon \bar{v} &\geq 0.
 \end{aligned}$$

Hence $(\bar{x}, \bar{w} + \epsilon \bar{v})$ is a Karush-Kuhn-Tucker point of (2) for any $\epsilon \geq 0$.

Case 2: $\bar{\gamma} > 0$. When $\bar{\gamma} > 0$, it follows from (4) that $(\bar{x}, \bar{u} = \bar{v}/\bar{\gamma})$ is a Karush-Kuhn-Tucker point of (2) with $\epsilon = \bar{\epsilon} = 1/\bar{\gamma}$. From (4) and (5) we have for $\bar{\gamma} > 0$ and $\lambda \in [0, 1]$ that

$$\begin{aligned}
 \frac{\lambda}{\bar{\gamma}} \nabla f(\bar{x}) - A^T \left((1-\lambda)\bar{w} + \lambda \frac{\bar{v}}{\bar{\gamma}} \right) + p &= 0, \\
 A\bar{x} &\geq b, \\
 \left((1-\lambda)\bar{w} + \lambda \frac{\bar{v}}{\bar{\gamma}} \right) (A\bar{x} - b) &= 0, \\
 (1-\lambda)\bar{w} + \lambda \frac{\bar{v}}{\bar{\gamma}} &\geq 0.
 \end{aligned}$$

Hence $(\bar{x}, (1-\lambda)\bar{w} + \lambda(\bar{v}/\bar{\gamma}))$ is a Karush-Kuhn-Tucker point for (2) for $\epsilon = \lambda\bar{\epsilon} = \lambda/\bar{\gamma}$ and $\lambda \in [0, 1]$.

The last statement of the theorem follows from the standard sufficiency theory of nonlinear programming [2, Thm. 10.1.2]. \square

We can apply the same proof technique above to a considerably more general problem than (1), namely to the nonlinear programming problem:

$$\text{Minimize } \theta(x) \quad \text{subject to } g(x) \leq 0, \quad h(x) = 0,
 \tag{6}$$

where θ, g and h are functions from R^n into R, R^m and R^k respectively. However because of a constraint qualification restriction the results apply to a narrow class outside linear programs. Hence we shall merely state the result and omit the proof

which is quite similar to the proof of Theorem 1. We shall again associate with (6) a perturbed problem, namely for some $\varepsilon \geq 0$

$$(7) \quad \text{Minimize } \theta(x) + \varepsilon f(x) \quad \text{subject to } g(x) \leq 0, \quad h(x) = 0,$$

where f is from R^n into R . We shall assume that (6) has a local solution at \tilde{x} with minimum value of $\bar{\theta} = \theta(\tilde{x})$ and that B is the open ball with center \tilde{x} such that $\theta(\tilde{x}) \leq \theta(x)$ for all x in B satisfying the constraints $g(x) \leq 0$ and $h(x) = 0$. We further admit the possibility of the nonuniqueness of \tilde{x} and define $\tilde{S} = \{x | \theta(x) = \bar{\theta}, g(x) \leq 0, h(x) = 0, x \in B\}$.

THEOREM 2. *Let (6) have a nonempty set \tilde{S} of local optimal solutions satisfying a constraint qualification. Let θ, g, h and f be differentiable on \tilde{S} and let the nonlinear program*

$$(8) \quad \begin{aligned} &\text{Minimize } f(x) \quad \text{subject to } g(x) \leq 0, \\ &h(x) = 0, \\ &\theta(x) \leq \bar{\theta}, \quad x \in B \end{aligned}$$

have a Karush–Kuhn–Tucker point $(\bar{x}, \bar{v}, \bar{s}, \bar{\gamma})$ in $R^{n+m+k+1}$. Then there exists an $\bar{\varepsilon} > 0$ such that for each ε in $[0, \bar{\varepsilon}]$ there exists a $(\bar{u}, \bar{r}) : [0, \bar{\varepsilon}] \rightarrow R^{m+k}$ such that $(\bar{x}, \bar{u}(\varepsilon), \bar{r}(\varepsilon))$ is a Karush–Kuhn–Tucker point for the perturbed problem (7), and \bar{x} is also a local solution of the nonlinear program (6). In fact it is possible to take $\bar{\varepsilon} = 1/\bar{\gamma}$ when $\bar{\gamma} > 0$ and $\bar{\varepsilon}$ as any positive number when $\bar{\gamma} = 0$.

The main cause of the restrictive nature of this theorem outside linear programming is that in order for (8) to have a Karush–Kuhn–Tucker point its constraints must in general satisfy a constraint qualification. This is difficult when g, h and θ are nonlinear because of the constraint $\theta(x) \leq \bar{\theta}$. However when h is linear and θ and g are pseudoconcave or concave at \bar{x} then a constraint qualification is automatically satisfied [2, Thm. 11.3.6]. This is a somewhat restrictive extension which does however include the case when (6) is a linear program.

The rest of the paper is devoted exclusively to the perturbation (2) of the linear program (1). We will first show that, under appropriate assumptions, some element \bar{x} of the solution set \bar{S} of (1) will be a local (global) solution of $P(\varepsilon)$ for all sufficiently small $\varepsilon \geq 0$. We will then show that under slightly stronger assumptions, each global solution of $P(\varepsilon)$ for sufficiently small $\varepsilon \geq 0$ is also a solution of (1). We begin by assuming that $\min_{x \in \bar{S}} f(x)$ has a local (global) solution \bar{x} , so that there exists an open ball B with center \bar{x} such that $\bar{x} \in \bar{S} \cap B$ is optimal for the problem

$$(9) \quad \text{Minimize } f(x) \quad \text{subject to } x \in \bar{S} \cap B.$$

The proof of the subsequent results depends crucially on establishing a minimum rate of increase of px in certain directions that lead “away” from \bar{S} . These directions are related to *projections* of points in S on \bar{S} . The projection of a point x on \bar{S} is denoted by $\mu(x)$ with $\mu(x) \in \bar{S}$ and

$$\|\mu(x) - x\| = \min_{\mu \in \bar{S}} \|\mu - x\|,$$

where $\|\cdot\|$ denotes the ∞ norm throughout this paper unless otherwise subscripted. We state now the key result which gives the desired lower bound on $p(x - \mu(x))$ and give the proof in the Appendix.

LEMMA 1. *There exists an $\alpha > 0$ such that*

$$p(x - \mu(x)) \geq \alpha \|x - \mu(x)\| \quad \text{for all } x \in S.$$

We shall also need the following Lipschitz property on the perturbation function f . There exist positive numbers δ and K such that

$$(10) \quad f(\mu(x)) - f(x) \leq K \|x - \mu(x)\| \quad \text{for } x \in S \text{ and } \|x - \mu(x)\| \leq \delta.$$

Note that it follows from the definition of $\mu(x)$ that $\|x - \mu(x)\| \leq \delta$ whenever $\|x - \bar{x}\| \leq \delta$. With the above concepts we establish our next principal result.

THEOREM 3. *Let \bar{x} be a local solution of $\min_{x \in S} f(x)$. Then, for sufficiently small $\varepsilon \geq 0$, \bar{x} is both a global solution of the linear program (1) and a local solution of the perturbed problem (2) provided that either of the two following conditions holds:*

- (a) *The Lipschitz property (10) holds.*
- (b) *f is convex on some open set containing \bar{x} .*

Proof. (a) Let (10) hold and let $B = B(\bar{x}, \bar{\delta}) = \{x \mid \|x - \bar{x}\| < \bar{\delta}\}$ where $\bar{\delta}$ is chosen such that $0 < \bar{\delta} \leq \delta$ and \bar{x} is an optimal solution of (9). Note that if $x \in B(\bar{x}, \bar{\delta})$, then

$$\|x - \mu(x)\| \leq \|x - \bar{x}\| < \bar{\delta} \leq \delta.$$

Hence by part (b) of Lemma A.3 of the Appendix we have upon noting the equality $p\bar{x} = p\mu(x)$ that

$$\begin{aligned} \varepsilon f(\bar{x}) + p\bar{x} &\leq \varepsilon f(x) + px \quad \text{for } x \in S \cap B\left(\bar{x}, \frac{\bar{\delta}}{2}\right) \\ &\text{and } \varepsilon \in \left[0, \frac{\alpha}{K}\right]. \end{aligned}$$

Hence \bar{x} solves (2) for $\varepsilon \in [0, \alpha/K]$ with the added constraint that $x \in B(\bar{x}, \bar{\delta}/2)$.

(b) Let f be convex on $B(\bar{x}, r)$ for some $r > 0$. By Theorem 10.4 of [5] f is Lipschitzian on any open ball $B(\bar{x}, \delta)$ with $\delta < r$, and again we have that

$$\|x - \mu(x)\| \leq \|x - \bar{x}\| < \delta \quad \text{for } x \in B(\bar{x}, \delta).$$

Hence because f is Lipschitzian on $B(\bar{x}, \delta)$, the first inequality of (10) holds for $x \in B(\bar{x}, \delta)$, and because f is convex on $B(\bar{x}, \delta)$, \bar{x} is an optimal solution of (9) with $B = B(\bar{x}, \delta)$. Again by part (b) of Lemma A.3 of the Appendix we have that

$$\begin{aligned} \varepsilon f(\bar{x}) + p\bar{x} &\leq \varepsilon f(x) + px \quad \text{for } x \in S \cap B\left(\bar{x}, \frac{\delta}{2}\right) \\ &\text{and } \varepsilon \in \left[0, \frac{\alpha}{K}\right]. \end{aligned}$$

Hence \bar{x} solves (2) for $\varepsilon \in [0, \alpha/K]$ with the added constraint that $x \in B(\bar{x}, \delta/2)$. \square

Example 1. To illustrate the need for the Lipschitz property (10), let $x \in \mathbb{R}^1$, let $S = \{x \geq 0\}$, $p = 1$, $f(x) = -x^{1/2}$. Note that f is continuous on S , but does not have the Lipschitz property in a neighborhood of $\bar{S} = \{0\}$. (Note also that f is convex on S , but cannot be extended to a finite convex function on \mathbb{R}^1 .) In this case it is easily verified that $\bar{S}(\varepsilon) = \varepsilon^2/4$ for $\varepsilon \geq 0$ and thus $\bar{S}(\varepsilon)$ never includes $\{0\}$ for any positive ε .

Note that in Theorem 3, the Lipschitz property (10) is needed *only* for those $x \in S$ that lie in some open neighborhood of \bar{x} , since only such points are involved in the statement of the theorem and its proof. On the other hand by using the full strength of (10) and under slightly stronger assumptions than those of Theorem 3 we can show that each global solution of the perturbed problem (2) for sufficiently small $\varepsilon \geq 0$ is also a solution of the linear program (1). In particular we have the following.

THEOREM 4. *Let $\bar{x} \in \bar{S}$ be a solution of $\min_{x \in \bar{S}} f(x)$ and let $px + \varepsilon^* f(x)$ be bounded from below on S for some $\varepsilon^* > 0$. Then $\bar{x} \in \bar{S}(\varepsilon) \subset \bar{S}$ for sufficiently small $\varepsilon \geq 0$ provided that any of the following conditions holds:*

- (a) *The Lipschitz property (10) holds.*
- (b) *f is convex on some open convex set containing S .*
- (c) *f has continuous first partial derivatives on some open set containing \bar{S} and \bar{S} is compact.*

Proof. We will first establish that $\bar{x} \in \bar{S}(\varepsilon) \subset \bar{S}$ for sufficiently small $\varepsilon \geq 0$ under hypothesis (a) by showing that for sufficiently small $\varepsilon \geq 0$

$$(11) \quad p\bar{x} + \varepsilon f(\bar{x}) < px + \varepsilon f(x) \quad \text{for } x \in S \setminus \bar{S}$$

and

$$(12) \quad p\bar{x} + \varepsilon f(\bar{x}) \leq px + \varepsilon f(x) \quad \text{for } x \in \bar{S}.$$

Inequality (12) holds because \bar{x} minimizes f on \bar{S} . To establish (11), let $x \in S \setminus \bar{S}$; thus $x \neq \mu(x)$, and consider the two following cases.

Case 1: $0 < \|\mu(x) - x\| \leq \delta$. The strict inequality (11) follows from part (a) of Lemma A.3 of the Appendix for $\varepsilon \in [0, \alpha/K)$ upon noting that $p\bar{x} = p\mu(x)$.

Case 2: $\|\mu(x) - x\| > \delta$. Let ν be such that $px + \varepsilon^* f(x) \geq \nu$ for $x \in S$, so that $f(x) \geq \nu/\varepsilon^* - px/\varepsilon^*$. By defining

$$q = -p/\varepsilon^* \quad \text{and} \quad \rho = -\nu/\varepsilon^* + f(\bar{x}) + p\bar{x}/\varepsilon^*$$

we have that

$$f(\bar{x}) - f(x) \leq q(\mu(x) - x) + \rho \quad \text{for } x \in S.$$

Because $\|\mu(x) - x\| > \delta$ it follows upon using the Hölder inequality that

$$\varepsilon(f(\bar{x}) - f(x))/\|x - \mu(x)\| \leq \varepsilon\|q\|_1 + \varepsilon\rho/\delta \quad \text{for } x \in S$$

and consequently for ε small enough, that is $\varepsilon \in [0, \alpha/(\|q\|_1 + \rho/\delta))$, the right hand side of the last inequality is less than α . Thus, for such ε

$$\begin{aligned} \varepsilon(f(\bar{x}) - f(x)) &< \alpha\|x - \mu(x)\| \\ &\leq px - p\bar{x} \quad \text{(by Lemma 1).} \end{aligned}$$

This establishes (11) for this second case also.

Now note that hypothesis (c) implies (a) and that hypothesis (b) also implies (a) in the case that \bar{S} is compact [5, Thm. 10.4], so that the proof will be completed by showing that the result holds under hypothesis (b) even when \bar{S} is not compact. Let

$$T = \{x \mid \|x - \bar{x}\| \leq k\},$$

where k is some positive number, let $S' = S \cap T$, and let $\bar{S}' = \bar{S} \cap T$. Note that S' is a compact polyhedral set and that \bar{S}' is the set of optimal solutions of $\min_{x \in S'} px$, so that the preceding arguments imply that there exists an $\varepsilon' > 0$ such that $\bar{x} \in \bar{S}'(\varepsilon) \subset \bar{S}'$ for $\varepsilon \in [0, \varepsilon']$ where $\bar{S}'(\varepsilon)$ denotes the solution set of $\min_{x \in S \cap T} px + \varepsilon f(x)$. Now suppose that for some $\varepsilon \in [0, \varepsilon']$, $\bar{S}(\varepsilon)$ contains a point $\bar{x} \notin \bar{S}$. By the convexity of $px + \varepsilon f(x)$, $\bar{x} \in \bar{S}'(\varepsilon)$ implies that $\bar{x} \in \bar{S}(\varepsilon)$ (since a local solution of $P(\varepsilon)$ must also be a global solution), and consequently by the convexity of $\bar{S}(\varepsilon)$ it follows that

$$x(\lambda) = (1 - \lambda)\bar{x} + \lambda\bar{x} \in \bar{S}(\varepsilon) \quad \text{for all } \lambda \in [0, 1].$$

However, for $\lambda \in (0, 1]$ we have that $x(\lambda) \notin \bar{S}$ and hence $x(\lambda) \notin \bar{S}'$. But for sufficiently

small $\lambda > 0$, $x(\lambda) \in \bar{S}'(\varepsilon) \subset \bar{S}'$, which is a contradiction. Thus $\bar{S}'(\varepsilon) \subset \bar{S}'$ for $\varepsilon \in [0, \varepsilon']$, and since $\bar{x} \in \bar{S}'(\varepsilon)$ the theorem is established under hypothesis (b). \square

In the terminology of point-to-set mappings the result $\bar{S}'(\varepsilon) \subset \bar{S}'$ of Theorem 4 for $\varepsilon \geq 0$ sufficiently small implies that the mapping $\bar{S}'(\varepsilon)$ is *upper semi-continuous at 0* in a strong sense. (Note that if $px + \varepsilon f(x)$ is *not* bounded from below for any $\varepsilon > 0$, then the inclusion $\bar{S}'(\varepsilon) \subset \bar{S}'$ holds trivially, since $\bar{S}'(\varepsilon) = \emptyset$ for all $\varepsilon > 0$.)

To see that the compactness of \bar{S} is necessary in hypothesis (c) of Theorem 4 we give below an example in which the conclusion of Theorem 4 fails when the compactness assumption of part (c) is dropped.

Example 2. Let $x \in \mathbb{R}^2$, $p = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $S = \{(x_1, x_2) | 1 \leq x_1, 0 \leq x_2 \leq 1\}$ and $f(x) = -x_1x_2 + x_1^3x_2^2$. Note that on S , $px + \varepsilon f(x) \geq \varepsilon(-x_1x_2 + (x_1x_2)^2) \geq -\varepsilon/4$. Moreover,

$$\bar{S} = \{(x_1, x_2) | x_1 \geq 1, x_2 = 0\}$$

and $f(x) = 0$ on \bar{S} , so that $px + \varepsilon f(x) = 0$ on \bar{S} for all $\varepsilon \geq 0$. However for $0 < \varepsilon \leq 2$, $x_1 = 2/\varepsilon$, $x_2 = \varepsilon^2/16$, we have that $(x_1, x_2) \in S$, $px + \varepsilon f(x) = -\varepsilon^2/32 < 0$, and hence no solution of $\min_{x \in \bar{S}} f(x)$ can be in $\bar{S}'(\varepsilon)$, the solution set of $\min_{x \in S} px + \varepsilon f(x)$. It can also be shown that $\bar{S}'(\varepsilon)$ is nonempty for all $\varepsilon > 0$ so that $\bar{S}'(\varepsilon)$ is not contained in \bar{S} .

Note that in the case that the linear program (1) has a *unique* solution, many of the results above may be simplified. In particular, Theorems 1 and 4 yield the following. (See also Remark 4 in [3].)

COROLLARY 1. *Let \bar{S} consist of a single point \bar{x} . If f is differentiable at \bar{x} , then \bar{x} is a Karush-Kuhn-Tucker point of (2) for all sufficiently small $\varepsilon \geq 0$. If, in addition, $\bar{S}(\varepsilon^*) \neq \emptyset$ for some $\varepsilon^* > 0$, then $\bar{S}'(\varepsilon) = \{\bar{x}\}$ for all sufficiently small $\varepsilon \geq 0$.*

Proof. The first conclusion follows directly from Theorem 1. The second follows from the fact that $\bar{S} = \{\bar{x}\}$ implies that $\mu(x) = \bar{x}$ for all $x \in S$, so that the Lipschitz property (10) holds as a consequence of differentiability of f at \bar{x} . This part of the corollary then follows from Theorem 4. \square

A similar result also holds without assuming uniqueness in (1) if a strict convexity property is assumed instead.

COROLLARY 2. *If f is strictly convex on some open set containing S and if \bar{x} is the solution of $\min_{x \in \bar{S}} f(x)$, then $\bar{S}'(\varepsilon) = \{\bar{x}\}$ for all sufficiently small $\varepsilon > 0$.*

Proof. The proof follows from Theorem 3 and the fact that, for $\varepsilon > 0$, $px + \varepsilon f(x)$ is strictly convex and therefore assumes its minimum at not more than one point in S . \square

Appendix.

LEMMA A.1. *There exists an $\alpha > 0$ such that*

$$p(x - \mu(x)) \geq \alpha \|x - \mu(x)\| \quad \text{for all } x \in S.$$

Proof. Obviously the lemma holds trivially when $x \in \bar{S}$ or equivalently when $x = \mu(x)$. Suppose now that $x \in S \setminus \bar{S}$ and let e be a vector of ones in \mathbb{R}^n . Then

$$\begin{aligned} 0 < \|x - \mu(x)\| &= \text{Minimum}_{\delta, \mu} \{ \delta | -\delta e \leq \mu - x \leq \delta e, A\mu \geq b, p\mu \leq \bar{\theta} \} \\ &= \text{Maximum}_{y, v, \zeta, w} \{ x(y - v) - \bar{\theta}\zeta + bw | y - v - p\zeta + A^T w = 0, \\ &\quad ey + ev = 1, \quad y, v, \zeta, w \geq 0 \} \end{aligned}$$

(by linear programming duality)

$$\begin{aligned}
 \text{(A.1)} \quad &= \underset{y, v, \zeta, w}{\text{Maximum}} \{ \zeta(px - \bar{\theta}) + w(b - Ax) | y - v - p\zeta + A^T w = 0, \\
 & \quad \quad \quad ey + ev = 1, \quad y, v, \zeta, w \geq 0 \}. \\
 &= \zeta(x)(px - p\mu(x)) + w(x)(b - Ax) \\
 & \quad \quad \quad (\text{since } \bar{\theta} = p\mu(x) \text{ and } (y(x), v(x), \zeta(x), w(x)) \text{ is a} \\
 & \quad \quad \quad \text{solution to the maximum problem}) \\
 &\leq \zeta(x)p(x - \mu(x)) \\
 & \quad \quad \quad (\text{since } w(x) \geq 0, \text{ and } b - Ax \leq 0).
 \end{aligned}$$

Thus $\zeta(x) > 0$ for $x \in S \setminus \bar{S}$, and in addition

$$\text{(A.2)} \quad 1 \leq \zeta(x) \frac{p(x - \mu(x))}{\|x - \mu(x)\|} \quad \text{for } x \in S \setminus \bar{S}.$$

But since $\zeta(x)$ may be chosen as a component of a solution vertex of the linear program (A.1) and since the feasible region of (A.1) is independent of x and has a finite number of vertices, $\zeta(x)$ for $x \in S \setminus \bar{S}$ may be bounded as follows

$$\begin{aligned}
 \zeta(x) \leq \frac{1}{\alpha} := \text{maximum} \{ \zeta | (y, v, \zeta, w) \text{ is a vertex of } y - v - p\zeta + A^T w = 0, \\
 \quad \quad \quad ey + ev = 1, \quad y, v, \zeta, w \geq 0 \}.
 \end{aligned}$$

This bound on $\zeta(x)$ together with (A.2) establishes the lemma. \square

LEMMA A.2. *If $\bar{x} \in \bar{S}$ and $x \in R^n$ then*

$$\|\mu(x) - \bar{x}\| \leq 2\|x - \bar{x}\|.$$

Proof.

$$\begin{aligned}
 \|\mu(x) - \bar{x}\| &\leq \|\mu(x) - x\| + \|x - \bar{x}\| \\
 &\leq \|\bar{x} - x\| + \|x - \bar{x}\| \quad (\text{since } \mu(x) \text{ is the projection of } x \text{ on } \bar{S}) \\
 &= 2\|x - \bar{x}\|. \quad \square
 \end{aligned}$$

LEMMA A.3. *Let the Lipschitz condition (10) hold, let $\bar{x} \in \bar{S} \cap B$ be a solution of $\min_{x \in \bar{S} \cap B} f(x)$ with the ball $B = B(\bar{x}, \delta)$ for some $\delta > 0$. Then for $\|x - \mu(x)\| \leq \delta$ and $x \in S \cap B(\bar{x}, \delta/2)$*

$$\text{(a) } \varepsilon(f(\bar{x}) - f(x)) < p(x - \mu(x)) \quad \text{for } x \neq \mu(x) \text{ and } \varepsilon \in [0, \alpha/K]$$

and

$$\text{(b) } \varepsilon(f(\bar{x}) - f(x)) \leq p(x - \mu(x)) \quad \text{for } \varepsilon \in [0, \alpha/K].$$

Proof. Let $x \in S \cap B(\bar{x}, \delta/2)$ and $\|x - \mu(x)\| \leq \delta$; then

$$\begin{aligned}
 \varepsilon(f(\bar{x}) - f(x)) &\leq \varepsilon(f(\mu(x)) - f(x)) \quad (\text{since by Lemma A.2 } \mu(x) \in \bar{S} \cap B(\bar{x}, \delta)) \\
 &\leq \varepsilon K \|\mu(x) - x\| \quad (\text{by (10) and } \|x - \mu(x)\| \leq \delta), \\
 &< \alpha \|\mu(x) - x\| \quad (\text{for } \varepsilon \in [0, \alpha/K] \text{ and } x \neq \mu(x)), \\
 &\leq p(x - \mu(x)) \quad (\text{by Lemma A.1}).
 \end{aligned}$$

This establishes part (a) of the lemma. Part (b) follows by changing the strict inequality in the above string of inequalities to an inequality for the case of $\varepsilon \in [0, \alpha/K]$. \square

REFERENCES

- [1] W. KARUSH, *Minima of functions of several variables with inequalities as side conditions*, Master of Science Dissertation, Dept. of Mathematics, Univ. of Chicago, December 1939.
- [2] O. L. MANGASARIAN, *Nonlinear Programming*, McGraw-Hill, New York, 1969.
- [3] ———, *Uniqueness of solution in linear programming*, *Linear Algebra and Appl.*, 25 (1979), pp. 151–162.
- [4] ———, *Iterative solution of linear programs*, Computer Sciences Tech. Rep. 327, Univ. of Wisconsin, Madison, in preparation.
- [5] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.