CONTINUITY OF MATHEMATICAL PROGRAMS AND LAGRANGE MULTIPLIERS

by

John Semple

Department of Mathematics and Statistics
McGill University, Montreal

August, 1985

Submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the Master of Science degree in Applied Mathematics.

ABSTRACT

This thesis studies the behaviour of mathematical models in finite-dimensional optimization. The models are considered as input-output systems where the input is a data vector (or parameter), and the output consists of the feasible set, the set of optimal solutions, the optimal value, and the Lagrange multipliers. In particular we obtain various conditions which guarantee continuity of the output.

SOMMAIRE

L'objet de cette thèse est l'étude de modèles mathématiques d'optimatisation en dimension finie. Nous considérons ces modèles comme des systèmes à entrées-sorties où les entrées sont des vecteurs de données (ou des paramètres) et les sorties consistent de l'ensemble des solutions acceptables, l'ensemble des solutions optimales, les valeurs optimales et les multiplications de Lagrange. Nous obtenons en particulier diverses conditions qui garantissent la continuité des sorties.

TABLE OF CONTENTS

Introduction

In this thesis we study the stability of mathematical programming models. We work within the framework of input optimization, e.g. [11], [13], [14]. Associated with each input is the feasible set, the set of optimal solutions, the optimal value, and sometimes the Lagrange multipliers. The latter four will be considered the output. A model is locally stable with respect to a fixed input vector θ^* and a region $S(\theta^*)$, if the output changes continuously for all sequences $\theta_n \to \theta^*$, $\theta_n \in S(\theta^*)$. The general form of the model is

$$\begin{aligned} &\min_{(x)} f^0(x,\theta) \\ &\text{s.t.} \\ &\quad f^k(x,\theta) \le 0, \qquad k \in \mathcal{P} \triangleq \{1,\cdots,m\}, \\ &\qquad \theta \in I \end{aligned}$$

where $f^0(x,\theta)$ is the objective function, the $f^k(x,\theta)$ $k \in \mathcal{P}$ are the constraint functions, $\mathcal{P}$ is a finite index set, I is a convex set in $\mathbb{R}^p$, and $x \in \mathbb{R}^n$. We will stipulate further conditions on the general model as needed. Some of the conditions include convexity of $f^k(x,\theta)$ in x for each fixed θ, pseudo-convexity of the constraint functions, and joint continuity of the functions in the variable (x,θ).

The thesis is divided into three chapters. In Chapter I we develop and study the notion of stability following the ideas of [11], [12], [13], and [14]. In particular, we introduce two new regions of stability. One of the main results is a new necessary condition for stability of convex models. In Chapter II we extend the notion of stability to include the Lagrange multipliers. We obtain conditions for upper semicontinuity of the Lagrange multipliers over a region of stability. The main results were recently published in [8]. Chapter III demonstrates that under suitable conditions for a simplified model we may obtain an explicit representation of the Lagrange multipliers which proves to be differentiable.

Chapter I Stability of Convex Models

1.1 Introduction to Stability

In this section we introduce basic ideas concerning the stability of convex mathematical programming models. The general model is of the form

$$\begin{aligned} &\operatorname*{Min}_{(x)} f^0(x,\theta) \\ &\text{s.t.} \\ &\quad f^k(x,\theta) \le 0 \qquad k \in \mathcal{P},\ \theta \in I,\ x \in \mathbb{R}^n, \end{aligned}$$

where the $f^k(x,\theta)$ are jointly continuous in (x,θ) and convex in x for fixed θ, $k \in \mathcal{P} \cup \{0\}$. The parameter vector (or data input), θ, is confined to some convex set I, $I \subseteq \mathbb{R}^p$. We will study the behaviour of such models when the parameter is perturbed in some neighbourhood of a fixed $\theta = \theta^*$. In particular, we will obtain certain sets of parameters for which the optimal solutions and optimal value of (P,θ) change "continuously" as functions of θ. These sets will have a reference point (usually denoted by $\theta = \theta^*$) which determines the present state of the model.

Suppose the model is running with the parameter $\theta = \theta^*$; then we have the following "output":

$F(\theta^*) = \{x \in \mathbf{R}^n : f^k(x,\theta) \leq 0 \quad k \in P\}$ — **the feasible set**

$\tilde{x}(\theta^*)$ — an **optimal solution**

$\tilde{F}(\theta^*)$ — the set of **all optimal solutions**

$\tilde{f}(\theta^*)$ — the **optimal value** (i.e. $f^0(\tilde{x}(\theta^*))$).

With these we are able to define the concept of a **stable region** which was introduced and studied previously in, among others, [12] and [14].

Definition. The model (P,θ) is stable in a region $S \subseteq \mathbf{R}^p$ at $\theta^* \in S$ if, for some neighbourhood $N(\theta^*)$ of θ^*, both

(i) $\theta \in N(\theta^*) \cap S \Longrightarrow \tilde{F}(\theta) \neq \emptyset$ and

(ii) $\theta \in N(\theta^*) \cap S$ and $\theta \to \theta^*$ imply that the set $\tilde{F}(\theta)$ is bounded and all its accumulation points are in $\tilde{F}(\theta^*)$. ■

With the proviso that $\tilde{F}(\theta^*) \neq \emptyset$ and bounded, the set

$$M(\theta^*) = \{\theta \in \mathbf{R}^p : F(\theta^*) \subseteq F(\theta)\}$$

is a region of stability at θ^*. Two more regions of stability require the following definitions:

$\mathcal{P}^=(\theta) = \{k : f^k(x,\theta) = 0 \;\; \forall\, x \in F(\theta)\}$ — the **minimal index set of active constraints**

$\mathcal{P}^<(\theta) = \mathcal{P} \setminus \mathcal{P}^=(\theta)$

$F^=(\theta) = \{x : f^k(x,\theta) = 0 \quad k \in \mathcal{P}^=(\theta)\}$.

Again, provided $\tilde{F}(\theta^*)$ is nonempty and bounded, the sets

$$V(\theta^*) = \{\theta : F^=(\theta^*) \subseteq F^=(\theta) \text{ and } f^k(x,\theta) \le 0 \quad \forall\, x \in F(\theta^*) \quad k \in \mathcal{P}^=(\theta^*) \setminus \mathcal{P}^=(\theta)\}$$

$$W(\theta^*) = \{\theta : F^=(\theta^*) \subseteq F^=(\theta) \text{ and } \mathcal{P}^=(\theta) = \mathcal{P}^=(\theta^*)\}$$

are regions of stability at θ^*. These sets have been examined in [12]. Each of these sets share the property that they are independent of the objective function. However, each of these sets is restrictive in that they require either $F^=(\theta^*) \subseteq F^=(\theta)$ or $F(\theta^*) \subseteq F(\theta)$. As the next example indicates, it is possible to have $F^=(\theta^*) \cap F^=(\theta) = \emptyset$ for all $\theta \ne \theta^*$, yet the feasible set seems "well behaved".

Example 1.1. Consider the feasible set given by

$$f^1(x,y,\theta) = y - \theta^2 x \leq 0$$

$$f^2(x,y,\theta) = -y + \theta^2 x \leq 0$$

$$f^3(x,y,\theta) = -y - x + 1 \leq 0 \qquad \theta \in [0,\infty) \triangleq I$$

The feasible set is the one-dimensional ray

$$F(\theta)=\{(x,y) : y=\theta^2 x, \ x \geq \frac{1}{1+\theta^2}\}.$$

The minimal index set of active constraints is constant,

$$\mathcal{P}^{=}(\theta) = \{1,2\}$$

and so the set $F^{=}(\theta)$ is easily determined:

$$F^{=}(\theta) = \{(x,y) : y = \theta^2 x\}.$$

However, the regions of stability $V(\theta^*)$ and $W(\theta^*)$ are not useful here since both contain only a single point:

$$V(\theta^*) = W(\theta^*) = \{\theta^*\}$$

for each $\theta^* \in I$ (note that $F^=(\theta^*) \not\subseteq F^=(\theta)$ for all $\theta \neq \theta^*$). Yet $F^=(\theta)$ satisfies the following property: for every point $(x_0, y_0) \in F^=(\theta^*)$ and every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\|\theta - \theta^*\| < \delta \Longrightarrow \exists \text{ a point } (x,y) \in F^=(\theta) \text{ satisfying}$$
$$\|(x,y) - (x_0,y_0)\| < \epsilon.$$

Thus every point in $F^=(\theta^*)$ can be approached by a sequence of points in $F^=(\theta)$ as $\theta \to \theta^*$. To find regions of stability which take into account the behaviour exhibited by Example 1.1, we need some additional concepts. In particular, we need to define the concept of continuity for sets.

1.2 A New Region of Stability

As promised at the close of the last section, we will now develop regions of stability which are useful for cases such as Example 1.1. To do so we need the following definitions taken from [3].

Definition. (i) The point to set mapping $\Gamma: I \longrightarrow X$ is **lower semicontinuous** at $e_0 \in I$ if for each open set $\mathcal{A} \subseteq X$ satisfying $\mathcal{A} \cap \Gamma(e_0) \neq \emptyset$, there exists a neighbourhood $N(e_0)$ of e_0 such that for each e in $N(e_0)$, $\Gamma(e) \cap \mathcal{A} \neq \emptyset$.

(ii) The point to set mapping $\Gamma: I \longrightarrow X$ is **upper semicontinuous** at $e_0 \in I$ if for each open set $\mathcal{A}$, $\mathcal{A} \subseteq X$ and $\mathcal{A} \supseteq \Gamma(e_0)$ there exists a neighbourhood $N(e_0)$ such that for each $e \in N(e_0)$, $\Gamma(e) \subseteq \mathcal{A}$. ■

We will say the point to set mapping is continuous at e^* if it is both upper semicontinuous and lower semicontinuous at e^*.

One may easily check that the point to set mapping $\Gamma: e \longmapsto F^=(e)$ corresponding to Example 1.1 is lower semicontinuous at each $e = e^*$, $e^* \in I = [0,\infty)$. Interestingly enough, it is upper semicontinuous nowhere on $[0,\infty)$. The definition of lower semicontinuity is unwieldy and so we prove the following property stated as a lemma.

Lemma 1.2. *A point to set mapping $\Gamma: I \longrightarrow X$ is lower semicontinuous at $e^* \in I$ if and only if, given any sequence $e_n \rightarrow e^*$ and $v^* \in \Gamma(e^*)$, there exists a sequence $v_n \in \Gamma(e_n)$ such* that $\nu_n \to \nu^*$.

Proof. First suppose that Γ is lower semicontinuous at e^* and we are given $\nu^* \in \Gamma(e^*)$ and a sequence $e_n \to e^*$. Let u_n be such that

(1.1) $\|u_n - \nu^*\| = \mathrm{Min}\{\|z - \nu^*\| : z \in \overline{\Gamma}(e_n)\}$, $u_n \in \overline{\Gamma(e_n)}$ and suppose u_n does not converge to ν^*. Then there exists $\epsilon > 0$ such that

$$\|u_n - \nu^*\| \geq \epsilon$$

for infinitely many n. The open ball $B_\epsilon(\nu^*)$ of radius ϵ cannot satisfy the criterion given in the definition of lower semicontinuity, for given any neighbourhood of e^* there exists a point e_n in this neighbourhood such that $B_\epsilon(\nu^*) \cap \Gamma(e_n) = \emptyset$ as insured by (1.1) and our choice of ϵ. By contradiction, we have proved one direction of the implication.

Now suppose that for all sequences $e_n \to e^*$ and points $\nu^* \in \Gamma(e^*)$ there exists a sequence of points $\nu_n \in \Gamma(e_n)$ such that $\nu_n \to \nu^*$, yet Γ is not lower semicontinuous at e^*. Then there exists $\mathcal{a}$ open such that $\mathcal{a} \cap \Gamma(e^*) \neq \emptyset$ and $\mathcal{a} \cap \Gamma(e) = \emptyset$ for e arbitrarily close to θ^*. Thus there exists a ball $B_\delta(\nu^*) \subseteq \Omega$ with center $\nu^* \in \Gamma(\theta^*)$ and radius $\delta > 0$, and a sequence $\{\theta_n\}$ such that $\theta_n \to \theta^*$ yet

$$\Gamma(\theta_n) \cap B_\delta(\nu^*) = \emptyset.$$

Hence

$$\|z_n - \nu^*\| \geq \epsilon \text{ for all } z_n \in \Gamma(\theta_n)$$

and there cannot exist a sequence $\{\nu_n\}_{n=1}^{\infty}$, $\nu_n \in \Gamma(\theta_n)$ such that $\nu_n \to \nu^*$, contradicting our original hypothesis. This proves the second half of the implication and the lemma. ■

The preceding lemma deals with the property cited in Example 1.1. However, even if the point to set mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ is lower semicontinuous at θ^*, the model need not be stable. The following example is evidence.

Example 1.3. Consider the convex mathematical program

$$\begin{aligned} &\text{Min } f^0(x,e) = x \\ &\text{s.t.} \\ &f^1(x,e) = -x-1-e \le 0 \\ &f^2(x,e) = -e^2 x \qquad e \in I \triangleq [-1,1] \end{aligned}$$

around $e^* = 0$. For $e > 0$

$$\mathcal{P}^=(e) = \emptyset \text{ and so } F^=(e) = \mathbf{R}.$$

For $e^* = 0$,

$$\mathcal{P}^=(e) = \{2\} \text{ and } F^=(e^*) = \mathbf{R}.$$

Thus the mapping $\Gamma: e \longmapsto F^=(e)$ is continuous at $e^* = 0$, yet the model is not stable. Observe that

$$F(e) = \begin{cases} [0,\infty) & e \neq 0 \\ [-1,\infty) & e = 0 \end{cases}$$

and

$$\tilde{F}(e) = \{\tilde{x}(e)\} = \begin{cases} 0 & \text{if } e \neq 0 \\ -1 & \text{if } e = 0 \end{cases}$$

Obviously the optimal value experiences a jump at $\theta^* = 0$ and the model is not stable. We can conclude that the continuity of the mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ at θ^* does not guarantee stability at θ^*. However, the next theorem will present one set of sufficient conditions. We need to consider the set

$$R(\theta^*) = \{\theta : \mathcal{P}^{=}(\theta) = \mathcal{P}(\theta^*)\}.$$

To facilitate the proof we need the following lemma.

Lemma 1.4. *Consider the mathematical model (P,θ). Suppose that the mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ is lower semicontinuous at θ^*. Then for any sequence $\theta_n \to \theta^*$, $\theta_n \in R(\theta^*)$, and for each fixed $x_0 \in F(\theta^*)$ there exists a sequence of points $v_n \in F(\theta_n)$ such that $v_n \to x_0$ (in other words, lower semicontinuity of the mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ at θ^* implies lower semicontinuity of the mapping $\theta \longmapsto F(\theta)$ at θ^* on $R(\theta^*)$).*

Proof. Let x_0 be given and let θ_n be an arbitrary sequence satisfying $\theta_n \in R(\theta^*)$ for all n and $\theta_n \to \theta^*$. We construct a sequence which satisfies the claim stated in the lemma. Let $y_n \in F(\theta_n)$ be such that

$$\|x_0 - y_n\| = \text{Min}\{\|z - x_0\| : z \in F(e_n)\}.$$

Note that the minimum exists since x_0 is a fixed point and $F(e_n)$ is closed ($F(e_n)$ need not be compact). Suppose y_n does not converge to x_0, then there exists an $\epsilon > 0$ and a subsequence, $y_{n(\ell)}$, such that

(2.1) $$\|y_{n(\ell)} - x_0\| \geq \epsilon \quad \forall \ell,\ y_{n(\ell)} \in F(e_{n(\ell)}).$$

Choose $\hat{x} \in F^{=}(e^*)$ such that both

$$f^k(\hat{x}, e^*) < 0 \quad \text{for } k \in \mathcal{P}^{<}(e^*), \text{ and } \|\hat{x} - x_0\| < \epsilon/2$$

simultaneously. By the joint continuity.of the functions, there exists a $\delta > 0$ such that

$$\|(\hat{x}, e^*) - (x, e)\| < \delta \Longrightarrow f^k(x, e) < 0 \text{ for } k \in \mathcal{P}^{<}(e^*).$$

By the lower semicontinuity at e^*, there exists a ball of radius w, $B_w(e^*)$ such that for $\hat{x} \in F^{=}(e^*)$ and each $e \in B_w(e^*)$ we may find $\overline{x}(e) \in F^{=}(e)$ satisfying

(2.2) $$\|\bar{x}(e) - \hat{x}\| < \mathrm{Min}\{\epsilon/2, \delta/2\}.$$

If we choose $L \in \mathbb{N}$ sufficiently large then we have

(2.3) $$\ell \geq L \Rightarrow \|e_{n(\ell)} - e^*\| < \mathrm{Min}\{\delta/2, w\}.$$

Since

$$\|(\bar{x}(e_{n(\ell)}), e_{n(\ell)}) - (\hat{x}, e^*)\| \leq \|\bar{x}(e_{n(\ell)}) - \hat{x}\| + \|e_{n(\ell)} - e^*\|$$

we may combine (2.2) and (2.3) to conclude

$$\|(\bar{x}(e_{n(\ell)}), e_{n(\ell)}) - (\hat{x}, e^*)\| < \delta \text{ for } \ell \geq L.$$

But δ was chosen so that

$$\|(\bar{x}(e_{n(\ell)}), e_{n(\ell)}) - (\hat{x}, e^*)\| < \delta \Rightarrow f^k(\bar{x}(e_{n(\ell)}), e_{n(\ell)}) < 0 \text{ for}$$
$$k \in \mathcal{P}^{<}(e^*)$$

and thus for $\ell \geq L$ the $\bar{x}(e_{n(\ell)}) \in F^{=}(e_{n(\ell)})$ are feasible points. Moreover, for $\ell \geq L$

$$\|\bar{x}(e_{n(\ell)}) - x_0\| \leq \|\bar{x}(e_{n(\ell)}) - \hat{x}\| + \|\hat{x} - x_0\| < \epsilon/2 + \epsilon/2 = \epsilon$$

which contradicts (2.1). Therefore, no such ϵ exists and we conclude that

$$y_n \in F(\theta_n) \text{ and } y_n \to x_0$$

which proves the lemma. ■

It may not be obvious where the assumption $\mathcal{P}^{=}(\theta) = \mathcal{P}^{=}(\theta^*)$ in the region $R(\theta^*)$ was used. The proof needs this assumption when constructing feasible points since we are able to write the feasible set as the intersection of two regions each with a constant index set, namely $F^{=}(\theta^*)$ and the region which we define as

$$F^{<}(\theta) \triangleq \{x : f^k(x,\theta) < 0 \quad k \in \mathcal{P}^{<}(\theta)\}.$$

Then

$$F(\theta) = F^{<}(\theta) \cap F^{=}(\theta).$$

Now we are in a position to prove Theorem 1.5.

Theorem 1.5. *Consider the convex model* (P,θ) *at* θ^*. *Suppose that* $\tilde{F}(\theta^*) \neq \emptyset$ *and bounded. If the mapping* $\Gamma : \theta \longmapsto F^=(\theta)$ *is lower semicontinuous at* θ^*, *then*

$$R(\theta^*) \triangleq \{\theta : \mathcal{P}^=(\theta) = \mathcal{P}^=(\theta^*)\}$$

is a region of stability at θ^*.

Proof. We must show that for some "restricted" neighbourhood of θ^* $N(\theta^*) \cap R(\theta^*)$ both:

(i) $\tilde{F}(\theta)$ is nonempty and bounded and

(ii) Given $\theta_n \in R(\theta^*)$, $\theta_n \rightarrow \theta^*$, we have that $\tilde{x}(\theta_n)$ is bounded and all its accumulation points are in $\tilde{F}(\theta^*)$.

We start by proving (i). First choose a particular optimal solution at θ^*, say $\tilde{x}_0(\theta^*)$. Since $\tilde{F}(\theta^*)$ is bounded (by assumption), we may enclose it in a closed ball $K \subseteq \mathbb{R}^n$ with surface ∂K, such that

$$\partial K \cap \tilde{F}(\theta^*) = \emptyset.$$

Suppose that (i) is not true. Then there exists a sequence of points $e_n \in R(e^*)$, $e_n \to e^*$, and a sequence of points $x(e_n)$ which satisfy one of two properties: either

(a) There exist optimal solutions $x(e_n)$ (at e_n) with $x(e_n) \cap K = \emptyset$ for all n; or

(b) $x(e_n) \cap K = \emptyset$ and $f^0(x(e_n), e_n) \leq f^0(z, e_n)$ for all $z \in K$, for all $n \in \mathbf{N}$

((a) in case the set of optimal solutions is unbounded, (b) in case there is no optimal solution).

In either case we may also find $v(e_n) \in F(e_n)$ such that $v(e_n) \to \tilde{x}_0(e^*)$ by lemma 1.4. Then either of (a) or (b) will imply

$$(2.4) \qquad f^0(v(e_n), e_n) \geq f^0(x(e_n), e_n) \text{ for all } n.$$

Since $\partial K \cap \tilde{F}(e^*) = \emptyset$, there exists N such that for all $n \geq N$

$$v(e_n) \in K.$$

By convexity, $\lambda v(\theta_n)+(1-\lambda)x(\theta_n)\in F(\theta_n)$ for all λ, $0\le\lambda\le 1$. For each $n\ge N$ we may choose λ_n such that

$$\lambda_n v(\theta_n)+(1-\lambda_n)x(\theta_n)\in\partial K,$$

i.e. lies on the surface of K. Equation (2.4) and the convexity of $f^0(x,\theta)$ then lead to the following conclusion:

$$f^0(\lambda_n v(\theta_n)+(1-\lambda)x(\theta_n),\theta_n)\le f^0(v(\theta_n),\theta_n) \tag{2.5}$$

for all n, $n\ge N$. The surface ∂K is compact and so there exists a convergent subsequence

$$\lambda_{n(\ell)}v(\theta_{n(\ell)})+(1-\lambda_{n(\ell)})x(\theta_{n(\ell)})\xrightarrow[\ell\to\infty]{} z_0\in\partial K.$$

The continuity of $f^0(x,\theta)$ yields

$$f^0(z_0,\theta^*)\le f^0(\tilde{x}_0(\theta^*),\theta^*). \tag{2.6}$$

Moreover, the continuity of the constraints guarantees that the limit point z_0 is a feasible point at θ^*. By construction, $\partial K\cap\tilde{F}(\theta^*)=\emptyset$. Since $z_0\in\partial K$, it cannot be an optimal solution. However, we have already deduced that z_0 is a feasible point and

(2.6) necessitates that it is an optimal solution as well. This yields the desired contradiction and (a) is proved.

To prove (b), we know that given any $\epsilon_n \to \epsilon^*$, $\epsilon_n \in R(\epsilon^*)$, $\tilde{x}(\epsilon_n)$ is bounded by part (a). Thus $\tilde{x}(\epsilon_n)$ has an accumulation point which we denote by x_0. Since x_0 is an accumulation point, there exists a subsequence of $\tilde{x}(\epsilon_n)$, $\tilde{x}(\epsilon_{n(\ell)})$ such that

$$\tilde{x}(\epsilon_{n(\ell)}) \to x_0 \text{ as } \ell \to \infty.$$

By the joint continuity of the constraints, x_0 must be a feasible point at ϵ^*. Letting $\tilde{x}_0(\epsilon^*)$ denote a particular optimal solution at ϵ^*, we then have .

$$(2.7) \qquad f^0(x_0, \epsilon^*) \geq f^0(\tilde{x}_0(\epsilon^*), \epsilon^*) = \tilde{f}(\epsilon^*).$$

Recalling Lemma 1.4 we have that there exits a sequence of feasible points satisfying both

$$\nu(\epsilon_{n(\ell)}) \in F(\epsilon_{n(\ell)}) \text{ and } \nu(\epsilon_{n(\ell)}) \to \tilde{x}_0(\epsilon^*).$$

But we must have

$$f^0(v(\theta_{n(\ell)}),\theta_{n(\ell)}) \geq f^0(\tilde{x}(\theta_{n(\ell)}),\theta_{n(\ell)}).$$

Taking the limit as ℓ tends to ∞ we get

$$(2.8) \qquad f^0(\tilde{x}_0,\theta^*) = f^0(\tilde{x}_0(\theta^*),\theta^*).$$

(2.7) and (2.8) imply that

$$f^0(x_0,\theta^*) = f^0(\tilde{x}_0(\theta^*),\theta^*).$$

This, taken with the fact that $x_0 \in F(\theta^*)$ (again by the continuity of the constraints in (x,θ)), yields the result $x_0 \in \tilde{F}(\theta^*)$ and completes the proof of (b) and the theorem. ■

Corollary 1.6. *If $\tilde{F}(\theta^*) \neq \emptyset$ and bounded for problem (P,θ) and Slater's condition holds at θ^* (i.e., there exists a point $\hat{x}$ such that $f^k(\hat{x},\theta^*) < 0$ $k \in \mathcal{P}$), then the model is stable in a neighbourhood of θ^*.*

Proof. $F^=(\theta^*) = F^=(\theta) = \mathbf{R}^n$ in some neighbourhood of θ^* (since $\mathcal{P}^=(\theta^*) = \emptyset$). Hence the mapping $\Gamma:\theta \longmapsto F^=(\theta)$ is lower semicontinuous at θ^*, and the joint continuity of the constraints implies that $\mathcal{P}^{=}(\theta) = \emptyset$ in a neighbourhood of θ^*. Thus Theorem 1.5 applies and the proof of the corollary is complete. ■

Corollary 1.7. *Suppose that for model* (P,θ), $\tilde{F}(\theta^*) \neq \emptyset$ *and bounded. Then the region*

$$W(\theta^*) = \{\theta : \mathcal{P}^{=}(\theta) = \mathcal{P}^{=}(\theta^*),\ F^{=}(\theta^*) \subseteq F^{=}(\theta)\}$$

is stable.

Proof. We need only observe that the condition

$$F^{=}(\theta^*) \subseteq F^{=}(\theta)$$

implies that the mapping $\Gamma : \theta \longmapsto F^{=}(\theta)$ (restricted to $W(\theta^*)$) is lower semicontinuous at θ^*. The hypotheses of Theorem 1.5 are satisfied (restricted to $W(\theta^*)$) and the corollary follows. ■

What is peculiar about the preceding theorem is that it doesn't make any use of upper semicontinuity. In general, even relatively simple mappings need not be upper semicontinuous. In particular, rotations in R^2 are not generally upper semicontinuous as the following example demonstrates.

Example 1.8. Consider the feasible set described by the following function:

$$f^1(x,y,\theta) = |y-\theta x| - (y-\theta x) \le 0 \quad \theta \in I \triangleq [1,\infty).$$

If $(y-\theta x) < 0$ then $f^1(x,y,\theta) > 0$ and so the point (x,y) is not feasible. If $(y-\theta x) \ge 0$ then $f^1(x,y,\theta) = 0$ and (x,y) is a feasible point. Thus the feasible set is

$$F(\theta) = \{(x,y) : y \ge \theta x\}.$$

At a fixed θ every feasible point (x_0,y_0) satisfies the condition $f^1(x_0,y_0,\theta) = 0$, so we have $\mathcal{P}^=(\theta) = \{1\}$ and

$$F^=(\theta) = F(\theta) = \{(x,y) : y \ge \theta x\}.$$

We now show that the mapping $\Gamma : \theta \longmapsto F^=(\theta)$ is not upper semicontinuous at any $\theta^* \in I$. To do so consider the open set

$$\Omega(\theta,\epsilon) \triangleq \text{int}(F^=(\theta) + \{(0,-\epsilon)\})$$

where $\text{int}(G)$ denotes the interior of the set G, "+" denotes the

Minkowski sum of two sets, i.e.

$$G + E \triangleq \{z : z = x + y \text{ for some } x \in G,\ y \in E\},$$

and ϵ is some specified scalar strictly greater than zero. Suppose we are given $\epsilon > 0$ and $\theta^* \in I$, then

$$Q(\theta^*, \epsilon) \supseteq F^=(\theta^*).$$

However, for any $\theta \neq \theta^*$ we can find an $\hat{x}$ such that

$$\hat{x} \in F^=(\theta) \text{ but } \hat{x} \notin Q(\theta^*, \epsilon).$$

Without loss of generality suppose $\theta > \theta^*$. Then choose $x_0 < 0$ such that

(2.9) $$x_0 \cdot (\theta - \theta^*) < -\epsilon.$$

The point $(x_0, \theta x_0) \in F^=(\theta)$ but $(x_0, \theta x_0) \notin Q(\theta^*, \epsilon)$. If it were then there would exist $(x_1, y_1) \in F^=(\theta^*)$ such that

$$(x_1, y_1) + (0, -\epsilon) = (x_0, \theta x_0).$$

But this implies $x_1 = x_0$. Since (x_1, y_1) must satisfy $y_1 \geq e^* x_1$, we have $y_1 \geq e^* x_1 = e^* x_0$ which in turn means $e^* x_0 - \epsilon \leq e x_0$. Regrouping yields

$$x_0 \cdot (e - e^*) \geq -\epsilon$$

contradicting (2.9). Thus the mapping is not upper semicontinuous at any $e^* \in I$. Again, as in Example 1.1, the mapping $\Gamma : e \longmapsto F^=(e)$ is lower semicontinuous at every $e^* \in I$. To substantiate this claim consider any $e^* \in I$ and any open set G satisfying $G \cap F^=(e^*) \neq \emptyset$. Because $F^=(e^*)$ is a half plane we have

$$G \cap F^=(e^*) \neq \emptyset \implies G \cap \operatorname{int}(F^=(e^*)) \neq \emptyset.$$

Since $G \cap \operatorname{int}(F^=(e^*))$ is open and nonempty, there exists an open ball of radius δ and center $(\hat{x}, \hat{y}) \neq (0,0)$, $B_\delta(\hat{x}, \hat{y})$, such that

$$B_\delta(\hat{x}, \hat{y}) \subseteq G \cap \operatorname{int}(F^=(e^*)).$$

This requires that

$$\hat{y} \geq e^* \hat{x} + \delta.$$

Obviously, if $|\theta - \theta^*| < \delta/|\hat{x}|$ for $\hat{x} \neq 0$, or $|\theta - \theta^*| < \infty$ for $\hat{x} = 0$, we have that the point $(\hat{x},\hat{y}) \in F^=(\theta)$ (because $\hat{y} \geq \theta\hat{x}$) and since $(\hat{x},\hat{y}) \in G$,

$$G \cap F^=(\theta) \neq \emptyset.$$

G was arbitrary and this completes the proof of the claim. ∎

The problem in the preceding example is that $F^=(\theta)$ is unbounded. As in Example 1.1, the set does not seem "badly" behaved and we want to characterize such simple sets. If we restrict ourselves to a bounded region, and if $P^=(\theta)$ is constant in a neighbourhood of θ^*, then it turns out that we get the upper semicontinuity at θ^* automatically. This result is stated explicitly in the following theorem.

Theorem 1.9. *Suppose that the constraints of problem (P,θ) are jointly continuous in (x,θ) and $P^=(\theta) = P^=(\theta^*)$ in a neighbourhood of θ^*. Then given any compact set $K \subseteq \mathbb{R}^n$ the mapping $\Gamma: \theta \longmapsto F^=(\theta) \cap K$ is upper semicontinuous at θ^*.*

Proof. Assume not. Then there exists Ω open, K compact (closed and bounded) such that

$$\mathcal{Q} \supseteq F^{=}(e^*) \cap K \text{ yet } \mathcal{Q} \not\supseteq F^{=}(e) \cap K$$

for e arbitrarily close to e^*. Thus there must exist sequences e_n and $x(e_n)$ such that

$$e_n \rightarrow e^*, \; x(e_n) \in \mathcal{Q}^c \cap F^{=}(e) \cap K$$

$$\text{and } x(e_n) \in F^{=}(e^*) \cap K \text{ for all } n,$$

(here $\mathcal{Q}^c$ denotes the complement of $\mathcal{Q}$). However, $x(e_n) \in F^{=}(e) \cap K$ and is therefore bounded. Thus $x(e_n)$ has a convergent subsequence and we may assume, after a relabelling, that

$$x(e_n) \rightarrow \hat{x} \text{ as } n \rightarrow \infty.$$

Since $\mathcal{Q}^c$ is closed, $\hat{x} \in \mathcal{Q}^c$. But

$$f^k(x(e_n), e_n) = 0 \text{ for } k \in \mathcal{P}(e_n) = \mathcal{P}(e^*).$$

By continuity of the constraints,

$$0 = f^k(x(e_n), e_n) \rightarrow f^k(\hat{x}, e^*) = 0 \text{ as } n \rightarrow \infty \text{ for } k \in \mathcal{P}^{=}(e^*).$$

This implies that $\hat{x} \in F^{=}(\theta^*) \cap K$ which is a contradiction (since then $\hat{x} \in \Omega$ as well). Hence no such Ω exists and in fact the mapping $\Gamma: \theta \longmapsto F^{=}(\theta) \cap K$ is upper semicontinuous at θ^* for each K compact. ■

1.3. A Necessary Condition For Stable Perturbations

The situation is considerably more complicated when the minimal index set of active constraints is not constant, particularly when we do not take into account the behaviour of the objective function. In general, the minimal index set of active constraints may decrease over a region and the model may be stable regardless of the objective function. An example of such a region would be $M(\theta^*)$ or $V(\theta^*)$. However, we open this section by showing that there do not exist any stable regions independent of the objective function which experience an increase in the minimal index set of active constraints. More precisely, given a path connected region satisfying the property $P^{=}(\theta) \supsetneqq P^{=}(\theta^*)$ in every neighbourhood of θ^*, one can always produce an objective function for which the model is unstable. This is summarized in the following theorem.

Theorem 1.10. *Suppose we are given the model* (P, θ) *and a path connected region* $S(\theta^*)$ *such that* $P^=(\theta) \not\subseteq P^=(\theta^*)$ *in every "restricted" neighbourhood* $N(\theta^*) \cap S(\theta^*)$. *Then there exists an objective function* $f^0(x, \theta)$ *for which the model is unstable.*

Proof. We may assume that $\mathcal{P}^=(\theta^*) \neq \mathcal{P}$ since the result is trivial in this case. We may also assume that $S(\theta^*)$ contains more than one point, for if not, the proof is again trivial. Since $S(\theta^*)$ is a path connected region with more than one point and since $\mathcal{P}^=(\theta^*) \neq \mathcal{P}$, there exists a sequence, $\theta_n \in S(\theta^*)$, a point $\hat{x} \in F(\theta^*)$ and an index ℓ satisfying the following:

$$\text{(i)} \quad \theta_n \to \theta^* \text{ and } f^\ell(x, \theta_n) = 0 \quad \forall x \in F(\theta_n)$$

and

$$\text{(ii)} \quad f^\ell(\hat{x}, \theta^*) < 0 \quad \hat{x} \in F(\theta^*).$$

Thus $\ell \in \mathcal{P}^<(\theta^*) \setminus \mathcal{P}^<(\theta_n)$ for all n. Now we claim that there exists a ball with radius $\epsilon > 0$ centered at $\hat{x}$ such that

$$B_\epsilon(\hat{x}) \cap F(\theta_n) = \emptyset$$

for all but a finite number of n. If not then for every $\epsilon > 0$ there exists feasible points in the ball $B_\epsilon(\hat{x})$ for infinitely many n. We can then construct a sequence of feasible points ν_n such that

$$\nu_n \rightarrow \hat{x} \text{ as } e_n \rightarrow e^*.$$

But (i) implies that $f^\ell(\nu_n, e_n) = 0$ for all n. The joint continuity of the constraints then implies that the limit point $\hat{x}$ satisfies

$$f^\ell(\hat{x}, e^*) = 0,$$

which is a contradiction. Hence there exists an $\epsilon > 0$ such that $B_\epsilon(\hat{x}) \cap F(e_n) = \emptyset$ for all but a finite number of n. This is essentially the end since we may now choose the objective function

$$f^0(x, e) = \|x - \hat{x}\|^2 \tag{3.10}$$

which is uniquely minimized at e^* by $\tilde{x}(e^*) = \hat{x}$. The model with the objective function (3.10) must experience a jump in the optimal value at θ^*. ∎

One should also note that the only assumption on $\hat{x}$ is that

$$f^k(\hat{x},\theta^*) < 0 \quad k \in \mathcal{P}^<(\theta^*).$$

In fact, the condition $\mathcal{P}^=(\theta) \not\subseteq \mathcal{P}^=(\theta^*)$ in a neighbourhood of θ^* indicates a high degree of instability since we cannot get near any feasible point which satisfies the condition

$$f^k(x,\theta^*) < 0 \quad k \in \mathcal{P}^<(\theta^*).$$

In the one dimensional case this indicates a virtual collapse of the feasible set outside of θ^*. Also note that the path connected assumption was used only to insure the existence of a sequence θ_n such that $\theta_n \to \theta^*$, with $\theta_n \neq \theta^*$ for an infinite number of n.

Another way of stating Theorem 1.10 is that if we have a region of stability $S(\theta^*)$ which is independent of the objective function, then a necessary consequence is the existence of a neighbourhood $N(\theta^*)$ such that

$$\mathcal{P}^=(\theta) \subseteq \mathcal{P}^=(\theta^*) \text{ for all } \theta \in N(\theta^*).$$

As we shall see in the next section, the requirement that the mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ be lower semicontinuous at θ^* also implies the condition $\mathcal{P}^{=}(\theta) \subseteq \mathcal{P}^{=}(\theta^*)$ in some neighbourhood of θ^*.

1.4. More on the Mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$

At this point one might suppose that the lower semicontinuity condition on the mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ might be a necessary condition for stability. This need not be the case even for regions which remain stable for all jointly continuous objective functions. The next example serves as proof.

Example 1.11. Consider the feasible set determined by

$$f^1(x) = -x \leq 0$$
$$f^2(x) = x - 1 \leq 0$$
$$f^3(x,\theta) = \begin{cases} 0 & \text{for} \quad |x| \leq 1 \\ \theta(|x| - 1) & \text{for} \quad |x| > 1 \end{cases}$$
$$\theta \in I \triangleq [0,1]$$

around $\theta^* = 0$. For all $\theta \in I$ we have

$$\mathcal{P}^{=}(\theta) = \{3\}$$

$$F(\theta) = \{x : 0 \leq x \leq 1\}$$

and so the model is stable for all jointly continuous objective functions. However

$$F^{=}(\theta^*) = \mathbb{R} \text{ and } F^{=}(\theta) = \{x : -1 \leq x \leq 1\}$$

for $\theta \neq \theta^*$, $\theta \in I$. The mapping $\Gamma : \theta \longmapsto F^{=}(\theta)$ is clearly not lower semicontinuous at $\theta^* = 0$. For instance, one may take $G \triangleq \{x : 2 < x < 4\}$ and observe that

$$G \cap F^{=}(\theta^*) = G \text{ yet } G \cap F^{=}(\theta) = \emptyset.$$

for all $\theta \neq \theta^*$, proving the claim. ■

The lower semicontinuity condition may not be a necessary condition for stability but it does provide some properties which we will soon exploit. One of these is stated in the next lemma.

Lemma 1.12. *For the model* (P, θ) *suppose the mapping* $\Gamma : \theta \longrightarrow F^{=}(\theta)$ *is lower semicontinuous at* θ^*. *Then* $\mathcal{P}^{<}(\theta) \supseteq \mathcal{P}^{<}(\theta^*)$ *for a sufficiently small neighbourhood of* θ^*.

Proof. We may construct a point $\hat{x}$ with the following property:

$$f^k(\hat{x},e^*) < 0 \quad k \in P^<(e^*) \text{ and } \hat{x} \in F(e^*).$$

By the joint continuity of the constraints there exists a $\delta_1 > 0$ such that

$$(4.1) \quad \|(x,e) - (\hat{x},e^*)\|_2 \leq \delta_1 \implies f^k(x,e) < 0 \quad k \in P^<(e^*).$$

By the lower semicontinuity condition at e^*, there exists a $\delta_2 > 0$ s.t.

$$(4.2) \quad \|e - e^*\|_2 < \delta_2 \implies \exists\, \bar{x}(e) \text{ s.t. } \bar{x}(e) \in F^=(e)$$

$$\text{and } \|\bar{x}(e) - \hat{x}\|_2 < \delta_1/2.$$

Thus for any e satisfying $\|e - e^*\|_2 \leq \text{Min}\{\delta_1/2, \delta_2\}$ we may find $\bar{x}(e)$ satisfying (4.2). Since

$$\|(\bar{x}(e),e) - (\hat{x},e^*)\|_2 \leq \|\bar{x}(e) - \hat{x}\|_2 + \|e - e^*\|_2 \leq \delta_1/2 + \delta_1/2 = \delta_1$$

we must have $f^k(\bar{x}(e),e) < 0 \quad k \in P^<(e^*)$ by (4.1). Finally, since $\bar{x}(e) \in F^=(e)$, $\bar{x}(e)$ is a feasible point, and the lemma is proved.
■

The preceding lemma tells us that the lower semicontinuity condition at θ^* is sufficient to guarantee the existence of a neighbourhood $N(\theta^*)$ such that $\mathcal{P}^=(\theta^*) \supseteq \mathcal{P}^=(\theta)$ or $\mathcal{P}^<(\theta) \supseteq \mathcal{P}^<(\theta^*)$ for $\theta \in N(\theta^*)$. We will use this property in connection with the set

$$V_1(\theta^*) = \{\theta \in \mathbb{R}^p : f^k(x,\theta) \leq 0 \quad \forall x \in F^=(\theta) \quad k \in \mathcal{P}^<(\theta) \setminus \mathcal{P}^<(\theta^*)\}.$$

The proof of the following theorem is in Appendix B.

Theorem 1.13. *Consider the model* $(\mathcal{P},\theta)$. *Suppose* $\tilde{F}(\theta^*) \neq \emptyset$ *and bounded and the mapping* $\Gamma : \theta \longmapsto F^=(\theta)$ *is lower semicontinuous at* θ^*. *Then the set*

$$V_1(\theta^*) = \{\theta \in \mathbb{R}^p : f^k(x,\theta) \leq 0 \quad \forall x \in F^=(\theta) \quad k \in \mathcal{P}^<(\theta) \setminus \mathcal{P}^<(\theta^*)\}$$

is a region of stability at θ^*. ■

We will return to the set $V_1(\theta^*)$ in the next chapter where it assumes more importance.

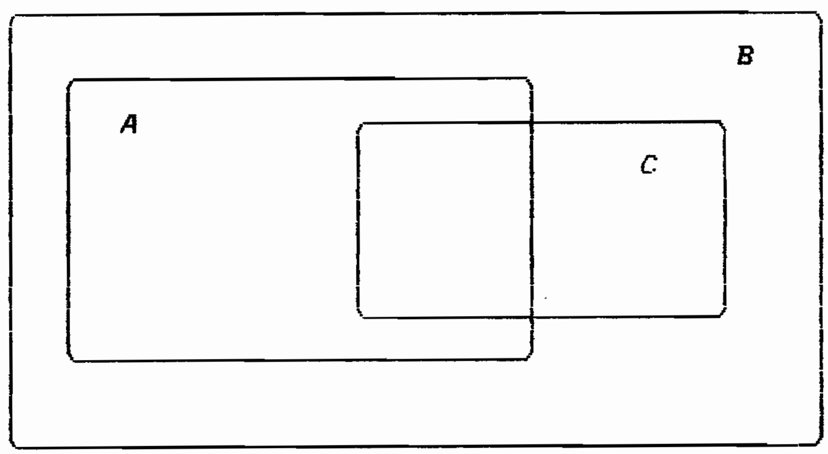
We close this chapter by organizing the results into a diagram. Let

A $\triangleq$ {models which are stable at $\theta = \theta^*$ independent of the objective function}

B $\triangleq$ {models for which $\wp^{<}(\theta) \supseteq \wp^{<}(\theta^*)$}

C $\triangleq$ {models for which the mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ is lower semicontinuous at θ^*}.

Then A, B and C share the following relationship:



Chapter 2. Continuity of the Lagrange Multipliers over Stable Regions

2.1 The Saddlepoint Inequality

In this section we investigate the continuity of the Lagrange multiplier function. The model we will work with is the same as that stated in Chapter 1:

$$(P,\theta) \qquad \begin{array}{ll} \text{Min} & f^0(x,\theta) \\ \text{s.t.} & \\ & f^k(x,\theta) \le 0 \quad k \in \mathcal{P} \triangleq \{1,\cdots,m\} \\ & \theta \in I \subseteq R^p, \end{array}$$

where $\theta = (\theta_i) \in R^p$ is the parameter vector, $x = (x_i) \in R^n$ is the vector variable; the functions $f^k(x,\theta)$ are jointly continuous in (x,θ) and convex in x for each fixed θ; $I \subseteq R^p$ is a convex set.

We express the optimality of a point for a fixed θ in terms of the restricted Lagrangian

$$(2.1) \qquad L^{<}(x,w;e) = f^{0}(x,e) + \sum_{k\in P^{<}(e)} w_k f^{k}(x,e)$$

We will denote the cardinality of the set $P^{<}(e)$ by $q(e)$, and the nonnegative octant of R^q by R^q_+. Then for a fixed e, a point $\tilde{x}$ is an optimal solution of (P,e) if, and only if, there exist nonnegative multipliers (u_i), $i \in P^{<}(e)$ which satisfy the restricted saddle point inequality

$$(2.2) \qquad L^{<}(\tilde{x},v,e) \leq L^{<}(\tilde{x},u,e) \leq L^{<}(x,u,e)$$

for every $v \in R^{q(e)}_+$ and every $x \in F^{=}(e)$. See, e.g. [13].

One might suppose that the saddle point inequality would hold for some generalized convex function. However, it turns out that (2.2) may not even hold for so-called pseudo-convex functions.

Definition. A differentiable function $f: R^n \longrightarrow R$ is pseudo-convex if

$$(2.3) \qquad \nabla f(x)\cdot(y-x) \geq 0 \implies f(y) \geq f(x).$$

The following example uses a convex objective function with a pseudo-convex constraint function.

Example 2.1. Consider the mathematical program

$$Min \quad f^0(x) = x$$

s.t.

$$g(x) \leq 0$$

$$g(x) = \begin{cases} -2x & \text{if} \quad x \leq 0 \\ -(x+1)^2 + 1 & \text{if} \quad 0 \leq x \leq 1 \\ x^2 - 6x + 2 & \text{if} \quad 1 \leq x \end{cases}$$

The function $g(x)$ has been pieced together so as to be differentiable on **R**. To verify that $g(x)$ is pseudo-convex we use (2.3). For $x \in (-\infty,3)$ $dg/dx(x) < 0$ and so

$$\frac{dg}{dx}(x)\cdot(y-x) \geq 0 \implies y \leq x.$$

Since g is monotonically decreasing on $(-\infty,3)$,

$$y \leq x \implies g(y) \geq g(x).$$

When combined with the earlier implication we have

$$\frac{dg}{dx}(x)\cdot(y-x) \geq 0 \implies g(y) \geq g(x)$$

and (2.3) holds on $(-\infty,3)$. On the interval $[3,\infty)$, $g(x) = x^2-6x+2$ and this function is convex, hence pseudo-convex. We will now show that there are no nonnegative multipliers which satisfy the saddle-point inequality (2.2). Suppose that they do exist. The optimal solution is $\tilde{x} = 0$ and so there exists $u_1 \geq 0$ such that

$$0 = \tilde{x} \leq x + u_1\cdot g(x) \text{ for all } x \in R.$$

For $x \leq 0$ we have

$$0 = x + u_1\cdot(-2x) \implies u_1 \geq 1/2.$$

But this is impossible. For example, when $\hat{x} = 3$, $g(\hat{x}) = -7$ and the saddle point inequality becomes

$$0 \leq 3 + u_1(-7) \implies u_1 < 3/7$$

which contradicts our earlier inequality, $u_1 \geq 1/2$. ■

Thus, for the remainder of this chapter we will focus on the problem (P,θ).

Since the Lagrange multipliers are not necessarily unique, we cannot talk about continuity in the normal sense. Furthermore, since inequality (2.2) depends on the index set $\mathcal{P}^{<}(\theta)$, we want to be sure that in any discussion of the Lagrange multipliers as functions of a parameter the set of indices we refer to remains fixed. This leads to the main result of this chapter, contained in the following section. ■

2.2 Continuity of the Lagrange Multiplier Function.

The closing remarks in the last section suggest that we must exercise caution in defining the Lagrange multiplier function. Indeed, we will insist that the mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ be lower semicontinuous at $\theta = \theta^*$ so as to insure that there exists a neighbourhood $N(\theta^*)$ of θ^* where $\mathcal{P}^{<}(\theta^*) \subseteq \mathcal{P}^{<}(\theta)$. Now we can make the following definition.

Definition. For the model (P,θ) suppose that the mapping

$\Gamma: \theta \longmapsto F^{=}(\theta)$ is lower semicontinuous at θ^*. Then we define the Lagrange multiplier function to be the point to set mapping

$$\mathcal{U}(\theta): \theta \longmapsto \{u_i(\theta) : i \in \mathcal{P}^{<}(\theta^*)\}. \blacksquare$$

Here $\{u_i(\theta) : i \in \mathcal{P}^{<}(\theta^*)\}$ is obtained by considering the set of all Lagrange multipliers at θ — $\{u_i(\theta) : i \in \mathcal{P}^{<}(\theta)\}$ — and looking at the $u_i(\theta)$ correspoinding to the index set $\mathcal{P}^{<}(\theta^*)$. We will use square brackets to indicate a particular Lagrange multiplier (or truncated Lagrange multiplier);

$$[u_i(\theta) : i \in \mathcal{P}^{<}(\theta^*)] \in \{u_i(\theta) : i \in \mathcal{P}^{<}(\theta^*)\}.$$

Note that $[u_i(\theta) : i \in \mathcal{P}^{<}(\theta^*)]$ is not necessarily a Lagrange multiplier at θ, however it will be if $\mathcal{P}^{<}(\theta) = \mathcal{P}^{<}(\theta^*)$. Thus the Lagrange multiplier function with respect to θ^* is obtained by taking the set of Lagrange multipliers at θ and removing those multipliers whose indices are in $\mathcal{P}^{=}(\theta^*)$. Note that the $u_i(\theta)$ are all defined for $i \in \mathcal{P}^{<}(\theta^*)$ in a neighbourhood of θ^* since $\mathcal{P}^{<}(\theta^*) \subseteq \mathcal{P}^{<}(\theta)$ by lemma 1.12.

We will consider the continuity properties of $\mathcal{U}(\theta)$ on several regions of stability. First we show that stability of the model (P,θ) does not guarantee continuity of the Lagrange multiplier function.

Example 2.2. Consider the model

$$\begin{aligned} &\text{Min } f^0(x) = x \\ &\text{s.t.} \\ &\qquad f^1(x,\theta) = -\theta^2 x \le 0 \\ &\qquad f^2(x,\theta) = -\theta^2 - x \le 0 \end{aligned}$$

near $\theta^* = 0$. Here

$$p^<(\theta) = \begin{cases} \{2\} & \text{if } \theta = 0 \\ \{1,2\} & \text{if } \theta \neq 0 \end{cases}$$

and $F^=(\theta) = \mathbb{R}$, for every θ (thus the mapping $\Gamma: \theta \longrightarrow F^=(\theta)$ is both upper semicontinuous and lower semicontinuous hence continuous). Also $F(\theta) = \mathbb{R}_+ = [0,\infty)$ and the optimal solution is $\tilde{x}(\theta) = 0$ for all θ.

The two regions of stability introduced in Chapter 1,

$M(\theta^*)$ and $V(\theta^*)$ are given by

$$M(\theta^*) = \{\theta : F(\theta^*) \subseteq F(\theta)\} = \mathbf{R}$$

$$V(\theta^*) = \{\theta : F^=(\theta^*) \subseteq F^=(\theta),\ -\theta^2 x \leq 0 \text{ for } x \geq 0)\} = \mathbf{R}.$$

For $\theta \neq 0$, the saddle point inequality (2.2) becomes

$$\tilde{x}(\theta) + u_1 \cdot (-\theta^2 \tilde{x}(\theta)) + u_2 \cdot (-\theta^2 - \tilde{x}(\theta))$$
$$\leq \tilde{x}(\theta) \leq x + \tilde{u}_1(\theta) \cdot (-\theta^2 x) + \tilde{u}_2(\theta) \cdot (-\theta^2 - x)$$

for every $u_1 \geq 0$, $u_2 \geq 0$ and every $x \in F^=(\theta) = \mathbf{R}$. These inequalities are uniquely satisfied by

$$\tilde{u}_1(\theta) = 1/\theta^2 \text{ and } \tilde{u}_2(\theta) = 0.$$

However, at $\theta^* = 0$, the saddle point inequality (2.2) becomes

$$\tilde{x}(\theta^*) - u_2 \cdot (\tilde{x}(\theta^*)) \leq \tilde{x}(\theta^*) \leq x - \tilde{u}_2(\theta^*) \cdot x$$

for every $u_2 \geq 0$ and every $x \in F^=(\theta^*) = \mathbf{R}$. This inequality is uniquely satisfied by

$$\tilde{u}_2(\theta^*) = 1.$$

The Lagrange multiplier function $\mathcal{U}(\theta)$ is then

$$\mathcal{U}(\theta) = \tilde{u}_2(\theta) = \begin{cases} 1 & \text{for} \quad \theta = \theta^* = 0 \\ 0 & \text{for} \quad \theta \neq \theta^*. \end{cases}$$

(One must remember that the function $\mathcal{U}(\theta)$ is defined according to the index set $\mathcal{P}^{<}(\theta^*) = \{2\}$). Thus $\mathcal{U}(\theta)$ need not be continuous even on a stable region. However, the Lagrange multiplier function is continuous on certain stable regions. The trick is to characterize those constraints which are in the index set $\mathcal{P}^{=}(\theta^*) \cap \mathcal{P}^{<}(\theta)$. If we assume that the mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ is lower semicontinuous at θ^*, then on the set

$$V_1(\theta^*) = \{\theta \in \mathrm{R}^p : f^K(x,\theta) \leq 0 \ \forall \ x \in \mathrm{F}^{=}(\theta) \ \ K \in \mathcal{P}^{<}(\theta) \cap \mathcal{P}^{=}(\theta^*)\}$$

we have continuity of $\mathcal{U}(\theta)$ in the sense of Theorem 2.4. Before we prove the theorem, we need the following lemma.

Lemma 2.3. *Suppose for the model* (P, e) *that the mapping* $\Gamma: e \longmapsto F^{=}(e)$ *is lower semicontinuous at* e^*. *Then given any* $\overline{x} \in F(e^*)$ *and any sequence* $e_n \rightarrow e^*$ *with* $e_n \in V_1(e^*)$ *for all* n, *there exists a sequence of feasible points* $\hat{x}(e_n) \in F(e_n)$ *such that* $\hat{x}(e_n) \rightarrow \overline{x}$ *as* $n \rightarrow \infty$.

Proof. Given $\overline{x} \in F(e^*)$ we may find $\overline{y} \in F(e^*)$ arbitrarily close to $\overline{x}$ such that

$$f^K(\overline{y}, e^*) < 0 \qquad K \in \mathcal{P}^{<}(e^*).$$

Since $\overline{y} \in F^{=}(e^*)$ and the mapping $\Gamma: e \longmapsto F^{=}(e)$ is lower semicontinuous at e^*, for $e_n \rightarrow e^*$ $e_n \in V_1(e^*)$ there exists $y(e_n) \in F^{=}(e_n)$ such that

$$y(e_n) \rightarrow \overline{y} \text{ as } n \rightarrow \infty$$

by lemma 1.2. Moreover, the lower semicontinuity condition guarantees the existence of a neighbourhood $N(e^*)$ of e^* such that

$$\mathcal{P}^{<}(e) \supseteq \mathcal{P}^{<}(e^*)$$

and so for n sufficiently large

$$\mathcal{P}^{<}(e_n) \supseteq \mathcal{P}^{<}(e^*).$$

After a relabelling we may assume that $\mathcal{P}^{<}(e_n) \supseteq \mathcal{P}^{<}(e^*)$ for all n. Now we look at the indices in $\mathcal{P}^{<}(e_n) \cap \mathcal{P}^{<}(e^*) = \mathcal{P}^{<}(e^*)$. Because $y(e_n)$ tends to $\bar{y}$ as e_n tends to e^* in the limit, the joint continuity of the constraints implies

$$f^K(y(e_n), e_n) < 0 \quad K \in \mathcal{P}^{<}(e^*)$$

for all but a finite number of n. But $y(e_n)$ is also in $F^{=}(e_n)$, and the condition

$$f^K(y(e_n), e_n) \leq 0 \quad K \in \mathcal{P}^{<}(e_n) \setminus \mathcal{P}^{<}(e^*)$$

for $e_n \in V_1(e^*)$, combined with the strict inequality $f^K(y(e_n), e_n) < 0 \quad K \in \mathcal{P}^{<}(e^*)$, leads to the conclusion

$$y(e_n) \in F(e_n)$$

for n sufficiently large. Thus there exists a sequence of feasible points $y(\theta_n) \in F(\theta_n)$ converging to $\bar{y}$ as θ_n converges to θ^*. Since $\bar{y}$ can be chosen arbitrarily close to $\bar{x}$ (due to convexity of the constraint functions), the same must be true for $\bar{x}$. This proves the lemma. ■

We note that the preceding lemma is critical in proving that $V_1(\theta^*)$ is a region of stability (see Appendix B). The proof of Theorem 2.4 follows.

Theorem 2.4. *Consider the model (P,θ) at some $\theta = \theta^*$ and suppose $\tilde{F}(\theta^*) \neq \emptyset$ and bounded and that the mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ is lower semicontinuous at θ^*. Let $\mathcal{U}(\theta) = \{u_K(\theta) \| K \in \mathcal{P}^{<}(\theta^*)\}$ be the Lagrange multiplier function. If $\theta^i \in V_1(\theta^*)$ and $\theta^i \to \theta^*$ then*

(i) *Any sequence drawn from $\{u_K(\theta^i) : K \in \mathcal{P}^{<}(\theta^*)\}$ is bounded for all sufficiently large i;*

(ii) *The set of limit points of $\{u_K(\theta) : K \in \mathcal{P}^{<}(\theta^*)\}$ as $\theta \to \theta^*$ is non-empty and every limit point is in $\{u_K(\theta^*) : K \in \mathcal{P}^{<}(\theta^*)\}$.*

Proof. (i) Let $\bar{x} \in F(\theta^*)$ be such that

(2.6) $f^K(\bar{x},\theta^*) < 0 \quad K \in \mathcal{P}^<(\theta^*)$

$f^K(\bar{x},\theta^*) = 0 \quad K \in \mathcal{P}^=(\theta^*)$

There exists $\epsilon > 0$ such that

(2.7) $f^K(\bar{x},\theta^*) \leq -\epsilon < 0 \quad K \in \mathcal{P}^<(\theta^*)$.

Given a sequence $\theta^i \to \theta^*$, $\theta^i \in V_1(\theta^*)$, we can construct a sequence of feasible points $\hat{x}(\theta^i) \in F(\theta^i)$ such that

(2.8) $$\hat{x}(\theta^i) \to \bar{x} \quad \text{as} \quad i \to \infty$$

by lemma 2.3. From the saddle point inequality (2.2) we have

(2.9) $$f^0(x,\theta^i) + \sum_{K \in \mathcal{P}^<(\theta^*)} u_K(\theta^i) f^K(x,\theta^i) + \sum_{K \in \mathcal{P}^=(\theta^*) \cap \mathcal{P}^<(\theta^i)} u_K(\theta^i) f^K(x,\theta^i) \geq \tilde{f}(\theta^i)$$

for every $x \in F^=(\theta^i)$. Now we will use (2.8) to show that any sequence from $\{u_K(\theta^i) : K \in \mathcal{P}^<(\theta^*)\}$, $i = 1, \cdots$ is bounded. If not, then for at least one index $\ell \in \mathcal{P}^<(\theta^*)$ and a subsequence $\theta^{i,j}$ of θ^i, there exists a subsequence

$$[\hat{u}_K(\theta^{i,j}) : K \in \mathcal{P}^{<}(\theta^*)]_{j=1}^{\infty} \in \{u_k(\theta^{i,j}) : K \in \mathcal{P}^{<}(\theta^*)\}_{j=1}^{\infty}$$

such that

(2.10) $$\hat{u}_{\ell}(\theta^{i,j}) \to +\infty \quad \text{as} \quad j \to \infty.$$

After a relabelling we may assume that

(2.11) $$\hat{u}_{\ell}(\theta^{i}) \to +\infty \quad \text{as} \quad i \to \infty.$$

By the definition of $V_1(\theta^*)$ and the fact that the Lagrange multipliers are nonnegative, we must have

$$\sum_{K \in \mathcal{P}^{=}(\theta^*) \cap \mathcal{P}^{<}(\theta^i)} \hat{u}_K(\theta^i) f^K(\hat{x}(\theta^i), \theta^i) \le 0.$$

The inequality (2.9) becomes

$$f^0(\hat{x}(\theta^i), \theta^i) + \sum_{K \in \mathcal{P}^{<}(\theta^*)} \hat{u}_k(\theta^i) f^k(\hat{x}(\theta^i), \theta^i) \ge \tilde{f}(\theta^i)$$

for all $x \in F^{=}(\theta^i)$. Recall by (2.8) that $\hat{x}(\theta^i) \to \bar{x}$ as $i \to \infty$. Using this, (2.11), and taking the limit as i tends to infinity yields

$$\tilde{f}(\theta^*) \leq \infty.$$

(Since $V_1(\theta^*)$ is a stable region we must have $\tilde{f}(\theta^i) \to \tilde{f}(\theta^*)$ as $i \to \infty$. See Appendix B.) This is absurd and proves that (2.11) must be false, hence (i) holds.

(ii) Let θ^i be an arbitrary sequence satisfying both

$$\theta^i \in V(\theta^*) \text{ for all } i \text{ and } \theta^i \to \theta^* \text{ as } i \to \infty.$$

Moreover, let $[u_k(\theta^i) : k \in \mathcal{P}^<(\theta^*)]_{i=1}^{\infty}$ be an arbitrary sequence of truncated Lagrange multipliers drawn from the sets $\{u_k(\theta^i) : k \in \mathcal{P}^<(\theta^*)\}_{i=1}^{\infty}$. Since each component, $u_k(\theta^i)$ is bounded by part (i), there exists a subsequence of θ^i for which each component converges. Thus the set of limit points of $\{u_k(\theta) : k \in \mathcal{P}^<(\theta^*)\}$ as $\theta \to \theta^*$ is nonempty.

Now let $\tilde{u}_k$ denote the kth component of a limit point of

$\{u_k(\theta^i) : k \in \mathcal{P}^{<}(\theta^*)\}$ as $\theta \to \theta^*$. Then there exists $\theta^i \to \theta^*$ and a sequence of truncated Lagrange multipliers $[u_k(\theta^i) : k \in \mathcal{P}^{<}(\theta^*)]_{i=1}^{\infty}$ such that

$$u_k(\theta^i) \xrightarrow[i \to \infty]{} \hat{u}_k \quad \text{for each } k \in \mathcal{P}^{<}(\theta^*).$$

Let $\tilde{x}(\theta^i)$ be optimal solutions chosen from $\tilde{F}(\theta^i)$. Since $V_1(\theta^*)$ is a stable region and $\theta^i \in V_1(\theta^*)$ for all i, we conclude that $\tilde{x}(\theta^i)$ is bounded and has a convergent subsequence with a limit point which we denote by $\tilde{x}(\theta^*)$. Again, since $V_1(\theta^*)$ is a stable region, $\tilde{x}(\theta^*) \in \tilde{F}(\theta^*)$ (hence our notation is consistent). Taking a subsequence of $u_k(\theta^i)$ for which $\tilde{x}(\theta^i)$ converges to $\tilde{x}(\theta^*)$, then taking the corresponding subsequence of $u_k(\theta^i)$ and relabelling this subsequence so as to have single superscripts, we may assume $\theta^i \to \theta^*$, $\theta^i \in V_1(\theta^*)$ and

$$u_k(\theta^i) \to \tilde{u}_k \text{ and } \tilde{x}(\theta^i) \to \tilde{x}(\theta^*) \text{ as } i \to \infty.$$

To complete the proof we must show that $[\tilde{u}_k : k \in \mathcal{P}^{<}(\theta^*)]$ is in $\{u_k(\theta^*) : k \in \mathcal{P}^{<}(\theta^*)\}$. We use the saddle point inequality (2.2) to obtain

$$(2.12) \qquad f^0(\tilde{x}(\theta^i),\theta^i) + \sum_{k \in \mathcal{P}^<(\theta^i)} \nu_k f^k(\tilde{x}(\theta^i),\theta^i) \le \tilde{f}(\theta^i)$$

for every $\nu_k \ge 0$, $k \in \mathcal{P}^<(\theta^i)$ and

$$(2.13) \qquad \tilde{f}(\theta^i) \le f^0(x,\theta^i) + \sum_{k \in \mathcal{P}^<(\theta^i)} u_k(\theta^i) f^k(x,\theta^i)$$

for every $x \in F^=(\theta^i)$. By specifying

$$\nu_k = 0, \quad k \in \mathcal{P}^<(\theta^i) \setminus \mathcal{P}^<(\theta^*)$$

in (2.12), we can change the summation index to $\mathcal{P}^<(\theta^*)$. In the limit as $\theta^i \to \theta^*$, (2.12) becomes

$$f^0(\tilde{x}(\theta^*),\theta^*) + \sum_{k \in \mathcal{P}^<(\theta^*)} \nu_k f^k(\tilde{x}(\theta^*),\theta^*) \le \tilde{f}(\theta^*)$$

for every $\nu_k \ge 0$ $k \in \mathcal{P}^<(\theta^*)$. This is the left hand side of the saddle point inequality (note that we need not concern ourselves with the set $F^=(\theta^*)$ yet). We now prove the other side of the inequality.

The term on the right hand side of the saddle point inequality given by (2.13) can be reduced. Since

$$f^k(x,\theta^i) \leq 0 \quad k \in \mathcal{P}^<(\theta^i)\setminus\mathcal{P}^<(\theta^*) \quad \forall\, x \in F^=(\theta^i)$$

(again by the definition of $V_1(\theta^*)$) we conclude

$$\sum_{k \in \mathcal{P}^<(\theta^i)\setminus\mathcal{P}^<(\theta^*)} u_k(\theta^i) f^k(x,\theta^i) \leq 0 \qquad \forall\, x \in F^=(\theta)$$

and so

$$\tilde{f}(\theta^i) \leq f^0(x,\theta^i) + \sum_{k \in \mathcal{P}^<(\theta^*)} u_k(\theta^i) f^k(x,\theta^i)$$

for every $x \in F^=(\theta^i)$. Finally, since every point in $F^=(\theta^*)$ is a limit point of points in $F^=(\theta^i)$ and since $\tilde{f}(\theta^i) \to \tilde{f}(\theta^*)$ as $i \to \infty$, we must have

$$\text{(2.14)} \qquad \tilde{f}(\theta^*) \leq f^0(x,\theta^*) + \sum_{k \in \mathcal{P}^<(\theta^*)} \tilde{u}_k f^k(x,\theta^*)$$

for every $x \in F^=(\theta^*)$. Recall $\tilde{u}_k$ denotes the limit point of $u_k(\theta^i)$, $k \in \mathcal{P}^<(\theta^*)$. This proves the right hand side of the saddle point inequality for the point $(\tilde{x}(\theta^*),\tilde{u})$. ■

One should note the importance of the condition $f^k(x,\theta) \leq 0$ $\forall x \in F^=(\theta)$ $k \in \mathcal{P}^<(\theta) \cap \mathcal{P}^=(\theta^*)$. In example (2.2) the Lagrange multipliers are bounded and so have a convergent subsequence. However, the lack of this additional condition results in the limit point not being a Lagrange multiplier at θ^*. We now prove a number of corollaries.

Corollary 2.5. *Consider the convex model* (P,θ) *at* θ^* *and suppose* $\tilde{F}(\theta^*) \neq \emptyset$ *and bounded. In addition suppose that the mapping* $\Gamma: \theta \longmapsto F^=(\theta)$ *is lower semicontinuous at* θ^*. *Then the Lagrange multiplier function* $\mathcal{U}(\theta) = \{u_k(\theta) ; k \in \mathcal{P}^<(\theta^*)\}$ *is upper semicontinuous on* $V_1(\theta^*)$.

Proof. If not then since $\mathcal{U}(\theta^*)$ is bounded, there exists an open set $A \supseteq \mathcal{U}(\theta^*)$ but $A \not\supseteq \mathcal{U}(\theta)$ in any neighbourhood of θ^* $N(\theta^*) \cap V_1(\theta^*)$. Thus there exists a sequence $\theta^i \in V_1(\theta^*)$, $\theta^i \to \theta^*$ and a sequence of Lagrange multipliers $u(\theta^i) \notin A$. The sequence $u(\theta^i)$ is bounded by the first part of Theorem (2.4) and so has a limit point which must be in A^c. Since $A^c \cap \mathcal{U}(\theta^*) = \emptyset$, Theorem

(2.4) is contradicted and the corollary is proved. ■

Corollary 2.6. *Consider the convex model (P,e) at some* $e = e^*$. *Assume that* $\tilde{F}(e^*) \neq \emptyset$ *and bounded and that Slater's condition holds at* $e = e^*$. *If the Lagrange multipliers are unique for all* e *in some neighbourhood of* e^*, *then* $\mathcal{U}(e)$ *is continuous in the usual sense at* e^*.

Proof. We first note that when Slater's condition holds, i.e.

$$\text{“}\exists\, \hat{x} \text{ such that } f^k(\hat{x},e^*) < 0 \quad k \in \mathcal{P}\text{”}$$

then $\mathcal{P}^=(e^*) = \emptyset$ and $F^=(e^*) = \mathbb{R}^n$. The Lagrangian $L^<(x,u,e)$ becomes the more familiar Lagrangian

$$L(x,u,e) = f^0(x,e) + \sum_{k\in\mathcal{P}} u_k f^k(x,e).$$

More importantly, $V_1(e^*)$ reduces to a neighbourhood $N(e^*)$ and the lower semicontinuity condition is satisfied trivially. Since the Lagrange multiplier function is unique we must have

$$u(e^i) \rightarrow u(e^*) \text{ as } i \rightarrow \infty$$

by Theorem (2.4). ■

The above result was previously obtained by Golstein (see [4]), Eremin and Astafiev [2].

One might suppose that the conditions on $V_1(\theta^*)$ would lead to some stronger properties of the Lagrange multiplier function $\mathcal{U}(\theta)$ around θ^*. In particular, the possibility of $\mathcal{U}(\theta)$ being lower semicontinuous might be studied. However, the next example proves that the conditions on $V_1(\theta^*)$ and on Theorem (2.4) do not guarantee the lower semicontinuity of $\mathcal{U}(\theta)$ at θ^*.

bf Example 2.7. Consider the convex model

$$\begin{aligned}
&\text{Min } f^0(x,\theta) = x \\
&\text{s.t.} \\
&f^1(x,\theta) = -x - \theta^2 \leq 0 \\
&\qquad f^1(x,\theta) = \begin{cases} x^2 & x \leq 0 \\ 0 & x \geq 0 \end{cases} \qquad \leq 0
\end{aligned}$$

around $\theta^* = 0$. For every $\theta \in R^1$, $F^=(\theta) = R_+$, $\mathcal{P}^<(\theta) = \{1\}$. Thus the mapping $\Gamma: \theta \longmapsto F^=(\theta)$ is lower semicontinuous at θ^*, and $V_1(\theta^*) = R$. For every θ we have $\tilde{x}(\theta) = 0$. For $\theta \neq 0$ the right hand side of the saddle point inequality is

$$0 = \tilde{x}(\theta) \leq x + u_1 \cdot (-x - \theta^2) \quad \forall x \in F^{=}(\theta) = \mathbf{R}_+$$

which is satisfied uniquely by $u_1 = 0$. However, at $\theta^* = 0$ the right hand side of the saddle point inequality becomes

$$0 = \tilde{x}(\theta^*) \leq x + u_1 \cdot (-x) \quad \forall x \in F^{=}(\theta^*) = \mathbf{R}_+$$

which is satisfied for $u_1 \in [0,1]$. The point $u_1 = 1$ at θ^* is not the limit point of any Lagrange multiplier as $\theta^i \to \theta^*$ ($\theta^i \neq \theta^*$). Therefore $\mathcal{U}(\theta)$ is not lower semicontinuous at θ^* (recall $\mathcal{U}(\theta)$ is defined with respect to θ^* as well). ■

2.3 Connection with Input Optimization.

In this section we establish the connection between the Lagrange multiplier function of Section 2.2 and the Lagrange multiplier function which exists for a locally optimal input. First we recall the idea of a locally optimal input from [12].

Definition. Consider the convex model (P,θ). An input $\theta^* \in S(\theta^*)$ is a **locally optimal input** over the region of stability $S(\theta^*)$ if both

(i) $\tilde{F}(\theta^*) \neq \emptyset$ and bounded, and

(ii) there exists a neighbourhood $N(\theta^*)$ of θ^* such that $\tilde{f}(\theta) \geq \tilde{f}(\theta^*)$ for every $\theta \in N(\theta^*) \cap S(\theta^*)$. ■

Characterizing an optimal input was done in [12] where it was required $F^=(\theta^*) \subseteq F^=(\theta)$. Under a suitable constraint qualification a necessary condition was found in terms of the Lagrangian

$$L_*^<(x,u,\theta) = f^0(x,\theta) + \sum_{K \in \mathcal{P}^<(\theta^*)} u_K f^K(x,\theta).$$

Note that this Lagrangian is different from the one defined at the beginning of this chapter since the summation is restricted to $\mathcal{P}^<(\theta^*)$. The theorem follows.

Theorem 2.8. *Consider the model* (P,θ). *Let θ^* be a locally optimal input over the region of stability $S(\theta^*)$, and let $x^* \in \tilde{F}(\theta^*)$. Then there exists a neighbourhood $N(\theta^*)$ of θ^* and a nonnegative vector function $\hat{U}$: $N(\theta^*) \cap S(\theta^*) \longrightarrow R_+^{q(\theta^*)}$ such that whenever $\theta \in N(\theta^*) \cap S(\theta^*)$,*

$$L_*^{<}(x^*,v,\theta^*) \le L_*^{<}(x^*,\hat{U}(\theta^*),\theta^*) \le L_*^{<}(x,\hat{U}(\theta),\theta)$$

for every $v \in R_+^{q(\theta^*)}$ *and every* $x \in F^{=}(\theta^*)$. ■

Since $F^{=}(\theta^*) \subseteq F^{=}(\theta)$ implies that the mapping $\Gamma: \theta \longmapsto F^{=}(\theta)$ is lower semicontinuous at θ^*, we can strengthen the statement of Theorem (2.8) considerably on the region of stability $W(\theta^*)$:

$$W(\theta^*) = \{\theta \in R^p : F^{=}(\theta^*) \subseteq F^{=}(\theta) \text{ and } \mathcal{P}^{=}(\theta) = \mathcal{P}(\theta^*)\}.$$

(This region was introduced in [14].) $W(\theta^*)$ is a subset of $V_1(\theta^*)$ and as one would expect, the nonnegative vector function in Theorem (2.8) can be identified as the Lagrange multiplier function of Theorem 2.4. The proof of this is trivial as one need only note that in the event $\mathcal{P}^{=}(\theta) = \mathcal{P}^{=}(\theta^*)$, the Lagrangian of theorem (2.8) and that of Theorem (2.4) coincide. As a consequence one may assume that the nonnegative vector function in Theorem (2.8) is continuous at θ^*.

Many examples point to the fact that near a locally optimal input there exist fixed nonnegative multipliers for which the right hand side of the inequality in Theorem 2.8 holds. That is, $U(\theta)$ could be taken to be constant in a sufficiently small neighbourhood of a locally optimal input θ^*. In fact, this turns out to be false and a counterexample is given in Appendix A.

Chapter 3. Continuity of the Lagrange Multipliers for Linear Models

3.1 An Explicit Representation of the Lagrange Multiplier Function

In this section the model is simplified and we assume that Slater's condition holds. Under a suitable hypothesis we will give an explicit representation of the Lagrange multiplier function and show that the function is continuously differentiable. We start by defining the Linear Model

$$(L,t) \qquad \min_{(x)} \langle a^0(t), x\rangle$$
$$\text{s.t.} \quad \langle a^j(t), x\rangle \le b^j(t) \quad j \in P \triangleq \{1,\cdots,m\} \qquad x \in R^n$$

where the $a^j(t)$, $j \in \wp \cup \{0\}$ are vector-valued functions of the scalar t whose i^{th} component is

$$\{a^j(t)\}_i \triangleq a_i^j(t) \triangleq \alpha_i^j t + \beta_i^j \quad \alpha_i^j,\ \beta_i^j \in R.$$

The $b_j(t)$ are differentiable functions of t (but not necessarily linear), and $\langle\ ,\ \rangle$ denotes the Euclidean inner product. For a fixed $t = t_0$ we have a linear program. Let

$B(x(t))$ denote the binding constraints at a feasible point $x(t)$. We begin by proving a simple lemma.

Lemma 3.1. *Consider the model (L,t) at some $t = t_0$. Suppose the optimal solution $\tilde{x}(t_0)$ is unique and Slater's condition holds. Then given any $\epsilon > 0$, there exists a $\delta > 0$ such that*

$$0 \leq |t - t_0| < \delta \Rightarrow \|\tilde{x}(t) - \tilde{x}(t_0)\| < \epsilon$$

for all optimal solutions at $\tilde{x}(t)$.

Proof. If not, then there exists $\epsilon > 0$ such that for some sequence $t_n \to t_0$, $\tilde{x}(t_n) \in \tilde{F}(t_n)$ we have

$$\|\tilde{x}(t) - \tilde{x}(t_0)\| \geq \epsilon. \tag{3.1}$$

However, the model is stable at t_0 since Slater's condition holds (see Chapter 1; Corollary 1.6). Thus $\tilde{x}(t_n)$ is bounded and has an accumulation point. Since the accumulation point must be in $\tilde{F}(t_0)$ we have

$$\tilde{x}(t_n) \to \tilde{x}(t_0) \text{ as } n \to \infty.$$

This contradicts (3.1) and the lemma is proved. ■

Note that we use the stability of the model implicitly in (3.1) by assuming that there exists an optimal solution to the perturbed model.

Now we are in a position to prove the main theorem of this chapter. First we recall that for differentiable convex mathematical programs, Slater's condition implies that the Kuhn-Tucker multipliers and Lagrange multipliers coincide.

Theorem 3.2. *Consider the model (L,t) at $t = t_0$. Suppose that the following conditions hold:*

(i) *$\tilde{x}(t_0)$ is unique*

(ii) *The $a^j(t_0)$ corresponding to the binding constraints are linearly independent, and that*

(iii) *Slater's condition holds.*

Then the Lagrange multiplier function is a differentiable path in $\mathbf{R}^q$ *($q = card\{B(\tilde{x}(t_0))\}$) in some neighbourhood of* t_0.

Proof. Without loss of generality we may assume that the set of binding constraints is

$$B(\tilde{x}(t_0)) = \{1, \cdots, k\} \quad k \leq m.$$

Consider those constraints which are non-binding, that is, those belonging to the index set $P \setminus B(\tilde{x}(t_0))$. Since the constraints are jointly continuous (this is not explicitly assumed but the claim is obvious from the form of (L,t)), there exists $\xi > 0$ such that

$$(3.2) \quad \|(y,t) - (\tilde{x}(t_0), t_0)\| < \xi \Rightarrow \langle a^j(t), y\rangle - b^j(t) < 0$$

$$j \in P \setminus B(\tilde{x}(t_0)).$$

By Lemma 3.1, the quantity $\|\tilde{x}(t_0) - \tilde{x}(t)\|$ can be made arbitrarily small for all $\tilde{x}(t) \in \tilde{F}(t)$ in a sufficiently small neighbourhood of t_0. Combining this with (3.2) implies the existence of a neighbourhood $N_1(t_0)$ such that

$$t \in N_1(t_0) \Rightarrow \langle a^j(t), \tilde{x}(t)\rangle - b^j(t) < 0 \quad j \in P \setminus B(\tilde{x}(t_0)).$$

and so for $t \in N_1(t_0)$,

$$B(\tilde{x}(t)) \subseteq B(\tilde{x}(t_0)).$$

Now define

$$f^j(x,t) \triangleq \langle a^j(t), x\rangle - b^j(t) \quad j \in \mathcal{P}.$$

The gradients taken with respect to x are

$$\nabla_x f^j(x,t) = a^j(t)$$

and depend only on t. Let $G(t)$ be the matrix of gradients corresponding to the binding constraints

$$G(t) = \begin{bmatrix} | & & | \\ \nabla_x f^1(x,t), & \cdots, & \nabla_x f^k(x,t) \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ a^1(t), & \cdots, & a^k(t) \\ | & & | \end{bmatrix}$$

By the hypothesis of the theorem $G(t_0)$ has full rank $(r(G(t_0))$ $= k = \text{card}\{B(\tilde{x}(t_0))\})$. By our knowledge of singular values (see, e.g. [9]) we know that for sufficiently small perturbations about $G(t_0)$, $G(t)$ will continue to have full rank. Since the elements in $G(t)$ are continuous functions of t, we are guaranteed that these perturbations can be made "small" in the Euclidean operator norm, defined for $A \in \mathbb{R}^{k \times n}$ as

$$\sup_{\|x\| = 1} \|Ax\|$$

where $\|x\|$ is the usual Euclidean norm $\|x\| = \langle x,x \rangle^{1/2}$. The preceding remarks mean that there exists a neighbourhood $N_2(t_0)$ such that

$$t \in N_2(t_0) \Rightarrow r(G(t)) = k.$$

Since Slater's condition holds, we turn to the Kuhn-Tucker equations for convex constraint functions and convex objective function (see [10]). Thus for every t in the smaller of the two neighbourhoods $N_1(t_0)$ and $N_2(t_0)$, the Kuhn-Tucker multipliers satisfy

(3.3) $$\left(G(t\ ) \right) u(t) = -\nabla_x f^0(t) = -a^0(t)$$

where $u(t) \geq 0$. Recall that the Lagrange multipliers and Kuhn-Tucker multipliers are identical for this model. One could differentiate the system (3.3) as it stands to obtain $u(t)$ (see [9]) but this form requires inverting matrices which depend on t. To get a better grip on the Lagrange multipliers $u(t)$ we rewrite (3.3) in the following form:

$$\left(G(t_0) + (G(t) - G(t_0)) \right) u(t) = -\nabla_x f^0(t) = -a^0(t).$$

$G(t_0)$ has full column rank and so there exists a $P \in \mathrm{R}^{n\times n}$ such that

$$PG(t_0) = \begin{pmatrix} I_k \\ \hline 0 \end{pmatrix}$$

where I_k denotes the $k\times k$ identity matrix. Hence

$$\text{(3.4)} \qquad \left(\begin{pmatrix} I_k \\ \hline 0 \end{pmatrix} + P[G(t) - G(t_0)] \right) u(t) = -P\cdot a^0(t).$$

But $G(t) - G(t_0)$ has a very specific form. Recall that the i^{th} component of $a^j(t)$ is $\alpha_i^j t + \beta_i^j$. Let $A \in \mathrm{R}^{n\times k}$ denote the matrix whose (i,j) entry is α_i^j. Then

$$G(t) - G(t_0) = (t - t_0)A.$$

Let us denote the first k rows of $P \cdot A$ by the matrix M. Furthermore, denote the top k rows of P by the matrix P_k. Truncating the last $(n-k)$ rows of (3.4) we get

$$[I_k + (t-t_0)M]u(t) = -P_k a^0(t). \tag{3.5}$$

Since M is a fixed matrix, $M \in \mathbf{R}^{k \times k}$, $\|M\| < \infty$ and so for $|t - t_0| < 1/\|M\|$ we have

$$|t - t_0|\,\|M\| < 1. \tag{3.6}$$

Here $\|M\|$ is the Euclidean operator norm. Thus the matrix on the left hand side of (3.5) is invertible and its inverse is given by

$$[I_k + (t - t_0)M]^{-1} = I_k + \Sigma_{j=1}^{\infty} (-1)^j (t - t_0)^j M^j. \tag{3.7}$$

We let the $S_n(t)$ denote the partial sum

$$S_n(t) \triangleq \sum_{j=0}^{n} (-1)^j (t - t_0)^j M^j$$

where $M^0 \triangleq I_k$. The matrix $S_n(t)$ can differentiated elementwise to obtain

(3.8) $$\frac{d}{dt}\left(S_n(t)\right) \triangleq S_n' = \sum_{j=1}^{n} j(t-t_0)^{j-1}(-1)^j M^j.$$

We may choose a neighbourhood $N_3(t_0)$ such that for $t \in \overline{N_3(t_0)}$ (3.6) holds (note it is valid for t in the closure of $N_3(t_0)$). Thus there exists a fixed $\alpha < 1$ such that

$$t \in N_3(t_0) \Longrightarrow |t-t_0| \, \|M\| \leq \alpha < 1.$$

We claim that both $S_n(t)$ and $S_n'(t)$ converge uniformly (using the Euclidean operator norm) on $\overline{N_3(t_0)}$. To show this we prove that the sequences are uniformly Cauchy. Given $\epsilon > 0$, choose $N \in \mathbf{N}$ such that both

$$\sum_{j=m}^{n} j\alpha^{j-1} < \epsilon$$

and

$$\sum_{j=m}^{n} \alpha^j < \epsilon$$

simultaneously for all m, $n \geq N$. For the same N we use the triangle inequality and the bounds on $|t-t_0|\|M\|$ to conclude

$$\|S_n(t) - S_m(t)\| \leq \sum_{j=m}^{n} \|(t-t_0)M\|^j \leq \sum_{j=m}^{n} \alpha^j < \epsilon$$

$$\|S_n'(t) - S_m'(t)\| \leq \sum_{j=m}^{n} j\|(t-t_0)M\|^j \leq \sum_{j=m}^{n} j\alpha^j < \epsilon$$

for all $t \in \overline{N_3(t_0)}$ and all $m, n \geq N$. Thus the sequences $S_n(t)$ and $S_n'(t)$ are Cauchy. Since the space of $k \times k$ matrices with the Euclidean operator norm form a Banach space, S_n' converges to a matrix in $R^{k \times k}$. However, we still do not know that this is the matrix of derivatives of (3.7). To verify that it is, we observe that

$$\|F\| \geq \max_{(i,j)} |F_{ij}|, \qquad F \in R^{k \times k}$$

and so the (i,j) element of $((t-t_0)M)^{\ell}$ satisfies

$$\|(t-t_0)M\|^{\ell} \geq \operatorname*{Max}_{(i,j)} |\{(t-t_0)^{\ell}M^{\ell}\}_{i,j}|.$$

Thus the elements of (3.8) converge uniformly, that is, the uniform convergence of $S_n'(t)$ in the Euclidean operator norm implies that the elements of $S_n'(t)$ converge uniformly as well. Since each element of $S_n(t)$ converges, we can conclude that

$$(3.9)\quad \lim_{n\to\infty} S_n'(t) = \frac{d}{dt}\left\{\lim_{n\to\infty} S_n\right\} = \frac{d}{dt}\left\{((I_K + (t-t_0)M)^{-1})\right\}.$$

For more on uniform convergence and differentiation of series, the reader is referred to [7].

From (3.7) and (3.4) we have

$$(3.10)\qquad u(t) = -\left(I_K + \sum_{j=1}^{\infty} (-1)^j (t-t_0)^j M^j\right) P_K a^0(t).$$

valid for t in the smallest of the three neighbourhoods, $N_1(t_0)$, $N_2(t_0)$, and $N_3(t_0)$.

Differentiating the product of a matrix and a vector, both functions of t, is similar to the usual produce rule from single variable calculus. For $A(t) \in R^{k \times k}$ and $v(t) \in R^k$ let

$$(3.11) \qquad \frac{d}{dt}\left\{A(t)\right\} \triangleq \dot{A}(t) \qquad\qquad \frac{d}{dt}\left\{v(t)\right\} = \dot{v}(t)$$

Then

$$\frac{d}{dt}\left\{A(t)\cdot v(t)\right\} = \dot{A}(t)v(t) + A(t)\dot{v}(t).$$

Recall that $a^0(t)$ is a vector-valued function whose i^{th} component is $\alpha_i^0 t + \beta_i^0$. Let

$$\alpha^0 \triangleq (\alpha_1^0, \cdots, \alpha_n^0)^t, \ \beta^0 \triangleq (\beta_i^0, \cdots, \beta_n^0)^t.$$

Then applying the differentiation result (3.11) to (3.10) we have

$$(3.12) \qquad \begin{aligned} \frac{d}{dt}\left\{u(t)\right\} = &-\left(\sum_{j=1}^{\infty} j(-1)^j (t-t_0)^{j-1} M^j\right) P_K (t\alpha^0+\beta^0) \\ &- \left(\sum_{j=0}^{\infty} (-1)^j (t-t_0)^j M^j\right) P_K \alpha^0 \end{aligned}$$

which is an explicit representation of the derivative of the Lagrange multipliers. The representation is valid for t in the smallest of the three neighbourhoods $N_1(t_0)$, $N_2(t_0)$ and $N_3(t_0)$. This completes the proof of the theorem. ■

Note that once t_0 is specified, the representations (3.10) and (3.12) depend only the quantities $(t - t_0)$, $P_k\alpha^0$, $.P_k\beta^0$, $M^jP_k\alpha^0$ and $M^jP_k\beta^0$. The latter four are constants and can be formed efficiently by recursion, $M^jP_k\alpha^0 = M(M^{j-1}P_k\alpha^0)$ and $M^jP_k\beta^0 = M(M^{j-1}P_k\beta^0)$, starting with the vectors $P_k\alpha^0$ and $P_k\beta^0$ respectively. For computations where it is desired to determine the Lagrange multipliers in a neighbourhood of t_0, this seems to be more efficient than inverting the matrix on the right-hand side of (3.3) repeatedly. The author is presently investigating the numerical possibilities associated with the representation (3.10) but the efficiency of using (3.10) and (3.12) compared to computing the Lagrange multipliers from scratch remains an open question.

Finally we draw the reader's attention to the fact that the representations in (3.10) and (3.12) incorporate perturbations in the objective function, left hand side and right hand side simultaneously. Since these representations are explicit in the variable t, one can track and study the behaviour of the

Lagrange multipliers in a neighbourhood of the present model ($t = t_0$).

Appendix A.

The example in this appendix disproves a conjecture concerning the nature of the nonnegative vector function associated with an optimal input. The model under consideration is the so called bi-convex model, a more restrictive model than (P,ϵ) of Chapter 1 since it also requires that the functions $f^k(x,\epsilon)$ $k \in \{0\} \cup \mathcal{P}$ be convex in ϵ for each fixed x. For this model, denoted by $P(x,\epsilon)$, the following conjecture was proposed.

"Let $P(x,\epsilon)$ be a perturbed bi-convex program. Let $\epsilon^* \in I$ be such that $\tilde{F}(\epsilon^*) \neq \emptyset$ and bounded with ϵ^* not being an extreme point of I. Suppose $S(\epsilon^*)$ is one of the three regions of stability, $M(\epsilon^*)$, $V(\epsilon^*)$, or $W(\epsilon^*)$ and suppose $S(\epsilon^*) \cap N(\epsilon^*) \cap I$ is convex. If ϵ^* is a locally optimal input over $S(\epsilon^*)$, then there exists a fixed vector u^* such that

A.1 $$L_*^{<}(x^*,u,\epsilon^*) \leq L_*^{<}(x^*,u^*,\epsilon^*) \leq L_*^{<}(x,u^*,\epsilon)$$

for $\epsilon \in S(\epsilon^*) \cap N(\epsilon^*) \cap I$, all $u \geq 0$, all $x \in F^{=}(\epsilon^*) \cap \Omega$, where $x^* \in \tilde{F}(\epsilon^*)$."

Here, Ω is a convex set $\Omega \subseteq R^n$. We will suppose below that the set Ω is R^n.

Although the conjecture is pretty, it is false, as the following example demonstrates.

Example A.1. Consider the biconvex model

$$\begin{array}{ll} \underset{(x)}{\text{Min}} & f^0(x)=|x+1|-1 \\ s.t. & g(x,\theta)=|\theta(x-1)|-\theta \leq 0 \\ & I=\{\theta : 1/2 \leq \theta \leq 2\} \end{array}$$

around the point $\theta^* = 1$. We leave it to the reader to convince himself that the model is bi-convex. Note that the feasible set is

$$F(\theta) = \{x : 0 \leq x \leq 2\}$$

for all $\theta \in I$, thus $M(\theta^*) = I$. Also note that every θ in I is a locally optimal input since $x = 0 \in F(\theta)$ for all $\theta \in I$ and $f^0 = |x+1| - 1$ is minimized at $x = 0$ over $F(\theta)$. In addition this minimum is unique (note that $f^0(x,\theta) > 0$ for $x > 0$). Hence, for all $\theta \in I$,

$$f^0(\tilde{x}(\theta)) = 0, \ \tilde{x}(\theta) = 0.$$

In particular, $\theta^* = 1$ is a locally optimal input, and θ^* is not an extreme point of I. It will turn out below that the Lagrange multipliers at θ^* are unique. For $\theta^* = 1$ the problem becomes

$$\underset{\substack{(x)\\ s.t.}}{\text{Min}} \quad f^0(x) = |x+1| - 1$$
$$|x-1| - 1 \leq 0$$

We seek $u^* \geq 0$ such that

$$\text{(A.2)} \qquad L(x^*, u, \theta^*) \leq L(x^*, u^*, \theta^*) \leq L(x, u^*, \theta^*)$$

(Slater's condition holds so we use the standard Lagrangian).

Since $\tilde{x}(\theta^*) = 0$ we have

$$g(\tilde{x}(\theta^*), \theta^*) = 0 \text{ and } f^0(\tilde{x}(\theta^*), \theta^*) = 0.$$

The saddle point inequality (A.2) becomes

$$0 + u\cdot(g(x^*, \theta^*)) \leq 0 + u^*\cdot 0 \leq f^0(x, \theta^*) + u^*\cdot g(x, \theta^*)$$

and then

$$0 + u\cdot(0) \leq 0 \leq |x+1| - 1 + u^*\cdot(|x-1| - 1)$$

which must hold for all $x \in R$ since $F^{=}(e^*) = R$, and all $u \geq 0$. The left inequality holds trivially. Choosing first $x \in (0,1)$ and then $x \in (-1,0)$ in the right inequality leads one to the conclusion

$$u^* = 1.$$

If there exists $u^* \geq 0$ such that (A.1) is to hold, it must be $u^* = 1$. Now we seek a neighbourhood $N(e^*)$ such that

$$\text{(A.3)} \qquad L(x^*,u,e^*) \leq L(x^*,u^*,e^*) \leq L(x,u^*,e)$$

holds for all $x \in F^{=}(e^*)$ and $u \geq 0$. Replacing the Lagrangian with the appropriate quantities reduces the right hand side of (A.3) to

$$\text{(A.4)} \qquad 0 = 0 + 1 \cdot 0 \leq |x+1| - 1 + 1 \cdot |e(x-1)| - e.$$

We will now show that (A.4) cannot hold for all $x \in F^{=}(e^*)$ and $u \geq 0$. In particular, for $x = 1/2$, $x \in F^{=}(e^*) = R$ we have

$$0 \leq 1/2 + |e/2| - e.$$

For $e > 1$ this becomes

$$0 \leq [1 - \theta]/2 < 0$$

which is a contradiction. Hence there does not exist a fixed $u^* \geq 0$ such that for the optimal input $\theta^* = 1$

$$L(x^*, u, \theta^*) \leq L(x^*, u^*, \theta^*) \leq L(x, u^*, \theta)$$

for some neighbourhood of θ^*, $N(\theta^*) \cap S(\theta^*) \cap I$, and all $x \in F^=(\theta^*)$, $u \geq 0$. ■

In conclusion, the nonnegative vector function associated with an optimal input is generally a non-trivial function of θ in some neighbourhood of θ^*.

Appendix B

In this appendix we prove that the set

$$V_1(e^*) \triangleq \{e : f^K(x,e) \leq 0 \; \forall x \in F^=(e) \quad K \in \mathcal{P}^<(e) \backslash \mathcal{P}(e^*)\}$$

is a region of stability whenever the mapping $\Gamma : e \longmapsto F^=(e)$ is lower semicontinuous at e^* and $\tilde{F}(e^*)$ is non-empty and bounded. The work for the proof has already been done and we only assemble the facts here. A review of Theorem 1.5 reveals that none of the properties of $R(e^*)$ are exploited. Instead, all that is needed is the property proved in Lemma 1.4:

"Given any $\bar{x} \in F(e^*)$ and a sequence $e_n \longrightarrow e^*$, there exists a sequence of feasible points $v_n \in F(e_n)$ such that $v_n \longrightarrow \bar{x}$ as $n \longrightarrow \infty$."

In other words, the mapping $e \longmapsto F(e)$ is lower semicontinuous at e^* whenever the mapping $\Gamma : e \longmapsto F^=(e)$ is lower semicontinuous at e^*.

The same property holds for $V_1(e^*)$ as demonstrated by Lemma 2.3 in Chapter 2. Thus the same proof given in Theorem 1.5 is valid for proving the stability of $V_1(\theta^*)$ except that we replace Lemma 1.4 with Lemma 2.3. We state this formally below.

Theorem 1.13. *Consider the model* (P,θ) *at some* $\theta = \theta^*$. *Suppose the mapping* $\Gamma: \theta \longmapsto F^{=}(\theta)$ *is lower semicontinuous at* θ^* *and that* $\tilde{F}(\theta^*)$ *is non-empty and bounded. Then the set* $V_1(\theta^*)$ *is a region of stability at* θ^*.

Proof. Replace lemma 1.4 with lemma 2.3 in the proof of Theorem 1.5. ■

REFERENCES

[1] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer, Non-Linear Parametric Optimization, Akademie-Verlag, Berlin, 1982.

[2] I. I. Eremin and N. N. Astafiev, Introduction to the Theory of Linear and Convex Programming, Nauka, Moscow, 1976 (in Russian).

[3] A. Fiacco, Introduction to Sensitivity and Stability Analysis in Non-Linear Programming, Academic Press, New York, 1983.

[4] E. G. Golstein, Konvexe Optimierung, WTB Mathematik-Physik, Akademie-Verlag, Berlin, 1975.

[5] A. Ben-Israel, A. Ben-Tal, and S. Zlobec, Optimality in Non-Linear Programming: A Feasible Directions Approach, Wiley-Interscience, New York, 1981.

[6] S. H. Jacobson, "Stability of bi-convex models in optimization", M.Sc. Thesis, Department of Mathematics, McGill University, Montreal, 1983.

[7] W. Rudin, Principles of Mathematical Analysis, McGraw-Hill, New York, 1976.

[8] J. Semple and S. Zlobec, "On the continuity of a Lagrangian multiplier function in input optimization", Mathematical Programming 33 (1985).

[9] G. W. Stewart, Matrix Computations, Academic Press, New York, 1973.

[10] S. Zlobec, Mathematical Programming (Lecture Notes), Department of Mathematics, McGill University, Montreal, 1977.

[11] S. Zlobec, "Input Optimization I: Optimal realizations of mathematical models", Mathematical Programming, 31 (1985), 245-268.

[12] S. Zlobec, R. Gardner, and A. Ben-Israel, "Regions of stability for arbitrarily perturbed convex programs", in: A. Fiacco, ed., Mathematical Programming with Data Perturbations, M. Dekker, New York, 1981, 69-89.

[13] S. Zlobec, "Characterizing an optimal input in perturbed

convex programming", Mathematical Programming 25 (1983), 109-121.

[14] S. Zlobec and A. Ben-Israel, "Perturbed convex programs: Continuity of optimal values", Operations Research Verfahren XXXI 1 (1979), 737-749.