CONSTRAINT QUALIFICATIONS IN INPUT OPTIMISATION

M. VAN ROOYEN[1], M. SEARS[1] AND S. ZLOBEC[2]

Abstract

We introduce assumptions in input optimisation that simplify the necessary conditions for an optimal input. These assumptions, in the context of nonlinear programming, give rise to conceptually new kinds of constraint qualifications.

1. Introduction

Constraint qualifications in nonlinear programming are conditions on the constraints that guarantee positiveness of the leading Lagrange multiplier (corresponding to the objective function) in the necessary conditions for optimality. Many papers have been written about constraint qualifications (abbreviation: CQ) particularly in the late sixties and in the seventies (see, for example, [9], [13], [14], [15], and [16]). Even in the cases when no such extraneous condition is needed, CQ may significantly simplify the theory and the numerical effort. For example, the BBZ complete characterisations of optimality in convex programming (see [2] and [3]) do not require any CQ. However, if a CQ, such as Slater's condition holds, the BBZ theory is significantly simplified and it recovers the classical KKT theory (of Karush [11] and Kuhn and Tucker [13]).

Recently a conceptually new level of optimisation, termed "input optimisation", has been formulated (see [23], [24], [25]). In input optimisation one is

[1]Department of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa 2001.
[2]Department of Mathematics and Statistics, McGill University, Burnside Hall, 805 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6.

concerned with mathematical *models* of the form

$$\text{Min}_{(x)} \, f^0(x, \theta)$$
$$\text{s.t.} \quad f^i(x, \theta) \leq 0, \quad i \in \mathscr{P} = \{1, \dots, m\} \qquad (P, \theta)$$
$$\theta \in I.$$

Here we assume that $f^j \colon R^n \times R^p \to R$ are continuous functions, $f^j(\cdot, \theta) \colon R^n \to R$ are convex, $j \in \{0\} \cup \mathscr{P}$, and $I \subset R^p$ is a convex set. Such models are termed *convex models*. (Note that for a fixed θ, the model (P, θ) becomes a usual mathematical *program*.) With every "input" (parameter) θ, we associate the "output" triple, that is

$$F(\theta) = \{x \colon f^i(x, \theta) \leq 0, \ i \in \mathscr{P}\},$$

the *feasible set*;

$$\tilde{F}(\theta) = \{\tilde{x}(\theta)\}$$

the *set of optimal solutions* $\tilde{x}(\theta)$, and

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

the *optimal value*. Suppose that the model is currently (initially) running with some input $\theta = \theta^0$. The main objective of input optimisation is to optimise the optimal value function $\tilde{f}(\theta)$, from $\theta = \theta^0$, by "stable" (that is, output continuous) perturbations of the input θ. The result is an *optimal input* θ^*, that is, an input θ^* with the property that

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

for every $\theta \in N(\theta^*) \cap S(\theta^*)$, where $N(\theta^*)$ and $S(\theta^*)$ are, respectively, some neighbourhood of θ^* and a region of stability at θ^* (see Definition 2.1 below). The corresponding program (P, θ^*) is termed an *optimal realisation* of the model (P, θ). Note that θ^* and (P, θ^*) depend on the initial θ^0, that is, determination of an optimal input is not generally a unique process.

Recently, a complete characterisation (that is, no CQ of any kind required) of an optimal input with respect to an arbitrary path-connected region of stability was given in [19] for convex models. The result, as expected, is complicated. Under the so-called "input constraint qualifications" (ICQ), the complete characterisation is significantly simplified and we recover the results from [22], and [25] (in Section 2). The main objective of this paper is to introduce a new kind of ICQ's, termed "modified input constraint qualifications" (MICQ). Under these conditions, we obtain yet another new and essentially different simplified form of necessary conditions for an optimal input (in Section 3). Then we show how ICQ's and MICQ's affect upper semicontinuity of various Lagrange multipliers, considered as point-to-set mappings (in Section 4), and we use these results to derive necessary conditions for an optimal input for differentiable functions (in

Section 5). The novelty is that our results are stated here for convex, rather than for more restrictive bi-convex, models. The results of this paper are readily applicable to the nonlinear (nonconvex) programs

$$\text{Min } f^0(z)$$
$$\text{s.t.} \qquad f^i(z) \le 0, \qquad i \in \mathscr{P}, \tag{P}$$

whenever, after identification $z = (x, \theta)$, the resulting "model" (P, θ) is convex. Since a locally optimal solution of (P) is necessarily a locally optimal input (but not vice versa!) we know that the necessary conditions for optimal inputs are also necessary conditions for optimal solutions. This means that we can state necessary conditions for a locally optimal solution of (P) without assuming any CQ! If an ICQ or an MICQ happens to be satisfied, then the necessary conditions for a locally optimal solution of (P) are simplified. The two conditions, ICQ and MICQ, are generally incomparable. They are a genuine product of input optimisation and a point-to-set approach to mathematical modelling and, as such, they do not seem to have counterparts in the usual theory of mathematical programming. We shall finish the paper by illustrating some of the above ideas on a well-known "pathological" program of Kuhn and Tucker. After an appropriate splitting $z = (x, \theta)$, the nonconvex program becomes a convex model to which all our results are easily applicable.

Finally, let us note that the results of this paper can be readily extended to the study of multi-objective mathematical programming models over regions of stability. One may use the approach recently suggested in [24]. However, it is not clear whether and how the results of this paper can be extended to the models with generalised convex functions.

2. Preliminaries

The notion of stability in input optimisation is centred around the point-to-set mapping $\Gamma : \theta \to F(\theta)$. Following [24] and [28] we say that a set $S \subset R^p$ is a *region of stability* at $\theta^* \in S$ if the point-to-set mapping Γ is lower semicontinuous at θ^* over S. This will be now formalised.

2.1 DEFINITION. Consider the convex model (P, θ) at some $\theta^* \in I$. A set $S \subset R^p$ is a region of stability at $\theta^* \in S$ if, for each open set $\mathscr{A} \subset R^n$ satisfying

$$\mathscr{A} \cap F(\theta^*) \ne \varnothing,$$

there exists a neighbourhood $N(\theta^*)$ of θ^* such that

$$\mathscr{A} \cap F(\theta) \ne \varnothing$$

for each $\theta \in N(\theta^*) \cap S$.

Lower semicontinuity of a point-to-set mapping is equivalent to its openness (see [1], [4] and [10]). On the other hand, the mapping Γ is always closed for our (P, θ), so Γ is continuous (in the sense of, say, Hogan [10]). This means that lower semicontinuity guarantees continuity. On the regions of stability the optimal solution mapping

$$\tilde{\Gamma} \colon \theta \to \tilde{F}(\theta)$$

is closed and the optimal value function $\tilde{f}(\theta)$ is continuous at θ^* for every *realistic* objective function. We recall the definition from [18] and [25].

2.2 DEFINITION. Consider the convex model (P, θ) at some $\theta^* \in I$. The objective function f^0 is realistic if $\tilde{F}(\theta^*) \neq \varnothing$ and bounded).

The complete characterisation of an optimal input θ^* is given in terms of the mapping

$$\gamma_* \colon \theta \to F_*^=(\theta),$$

where

$$F_*^=(\theta) = \{x \colon f^i(x, \theta) \leq 0, \ i \in \mathscr{P}^=(\theta^*)\}$$

and

$$\mathscr{P}^=(\theta) = \{i \in \mathscr{P} \colon x \in F(\theta) \Rightarrow f^i(x, \theta) = 0\}.$$

We shall also use the related set

$$F^=(\theta) = \{x \colon f^i(x, \theta) = 0, \ i \in \mathscr{P}^=(\theta)\}.$$

All our optimality conditions use the restricted Lagrangian, defined close to $\theta^* \in I$ by

$$L_*^<(x, u; \theta) = f^0(x, \theta) + \sum_{i \in \mathscr{P}^<(\theta^*)} u_i f^i(x, \theta), \tag{2.1}$$

where, for the sake of simplicity, we use the abbreviation

$$\mathscr{P}^<(\theta^*) = \mathscr{P} \setminus \mathscr{P}^=(\theta^*).$$

We shall also denote by $R_+^{q(\theta^*)}$ the non-negative orthant of the $q(\theta^*) = \operatorname{card} \mathscr{P}^<(\theta^*)$ dimensional Euclidean space.

There is still one small technical problem: in what follows, there will appear three different necessary conditions for optimality, each expressed in terms of the Lagrangian (2.1). However, the saddle-point inequalities will be restricted to different sets in the variable x, resulting in three essentially different Lagrange multiplier functions. For the lack of reasonable notation, and to make things as clear as possible, we shall replace "u" by "λ" in (2.1), but only in the characterisation of optimality, and use different u's in the two remaining necessary conditions.

We recall the characterisation from [19]:

2.3 THEOREM. *Consider the convex model* (P, θ) *with a realistic objective function at some* $\theta^* \in I$. *Let* $\tilde{x}(\theta^*)$ *be a corresponding optimal solution and let* S *be an arbitrary path-connected region of stability at* θ^*. *Then* θ^* *is a locally optimal input with respect to* S *if, and only if, there exists a neighbourhood* $N(\theta^*)$ *of* θ^* *and a non-negative vector function*

$$\Lambda: N(\theta^*) \cap S \to R_+^{q(\theta^*)}$$

such that, whenever $\theta \in N(\theta^*) \cap S$,

$$L_*^<(\tilde{x}(\theta^*), \lambda; \theta^*) \le L_*^<(\tilde{x}(\theta^*), \Lambda(\theta^*); \theta^*) \le L_*^<(x, \Lambda(\theta); \theta)$$

for every $\lambda \in R_+^{q(\theta^*)}$ *and every* $x \in F_*^=(\theta)$.

Examples exist showing that the above result does not generally hold at optimal inputs with respect to "unstable" regions.

The above result is not always easy to use to check optimality, so one may introduce an ICQ to simplify things. Following [22], and [25], an ICQ for the convex model (P, θ) at $\theta^* \in I$, with respect to a region of stability $S(\theta^*)$, is a condition on the constraints of the model with the property that for every $\theta \in N(\theta^*) \cap S(\theta^*)$, where $N(\theta^*)$ is a neighbourhood of θ^*, the system

$$
\begin{aligned}
f^0(x, \theta) &< \tilde{f}(\theta^*) \\
f^i(x, \theta) &< 0, \qquad i \in \mathscr{P}^<(\theta^*) \\
x &\in F^=(\theta^*)
\end{aligned}
\qquad (C, \theta)
$$

is inconsistent. For some regions of stability, such as

$$W(\theta^*) = \{\theta: F^=(\theta^*) \subset F^=(\theta), \ \mathscr{P}^=(\theta^*) = \mathscr{P}^=(\theta)\} \cap I$$

and

$$
\begin{aligned}
V_1(\theta^*) = \{\theta: F^=(\theta^*) \subset F^=(\theta) \text{ and } f^i(x, \theta) \le 0 \forall x \in F^=(\theta^*), \\
i \in \mathscr{P}^=(\theta^*) \setminus \mathscr{P}^=(\theta)\} \cap I
\end{aligned}
$$

it is easy to show that the system (C, θ) is always inconsistent at an optimal input θ^*. (The system (C, θ^*) is also inconsistent for every mathematical program (P, θ^*).) But for more general regions of stability, such as

$$Z(\theta^*) = \{\theta: F(\theta^*) \subset F^=(\theta) \text{ and } f^i(x, \theta) \le 0 \forall x \in F(\theta^*), \ i \in \mathscr{P}^=(\theta^*) \setminus \mathscr{P}^=(\theta)\},$$

the inconsistency of (C, θ) may not be easy to establish. In this case we use various "sufficient conditions" for ICQ such as

> "For every $\theta \in N(\theta^*) \cap Z(\theta^*)$, where $N(\theta^*)$ is a neighbourhood of θ^*, and for every $x \in F^=(\theta^*)$ such that
>
> $$f^i(x, \theta) < 0, \qquad i \in \mathscr{P}^<(\theta^*)$$
>
> it follows that
>
> $$f^i(x, \theta^*) \leq 0, \qquad i \in \mathscr{P}^<(\theta^*)".$$

This (nontrivial!) condition, called ICQ1 in the literature [22], is an ICQ for $Z(\theta^*)$. In particular the famous CQ of nonlinear programming:

> "There exists x^* such that $f^i(x^*, \theta^*) < 0,\ i \in \mathscr{P}$"

(the so-called Slater condition) is also an ICQ. Note that ICQ1 does not seem to have a known counterpart in usual nonlinear programming. If an ICQ holds, then the necessity part from Theorem 2.3 can be significantly simplified. We recall the result from [22]:

2.4 THEOREM. *Consider the convex model (P, θ) with a realistic objective function at some $\theta^* \in I$. Let $\tilde{x}(\theta^*)$ be a corresponding optimal solution. Suppose that θ^* is a locally optimal input with respect to a region of stability $S(\theta^*)$ and that an input constraint qualification is satisfied at θ^* with respect to $S(\theta^*)$. Then there exists a neighbourhood $N(\theta^*)$ of θ^* and a non-negative vector function*

$$\mathscr{U}: N(\theta^*) \cap S(\theta^*) \to R_+^{q(\theta^*)}$$

such that, whenever $\theta \in N(\theta^*) \cap S(\theta^*)$,

$$L_*^<(\tilde{x}(\theta^*), u; \theta^*) \leq L_*^<(\tilde{x}(\theta^*), \mathscr{U}(\theta^*); \theta^*) \leq L_*^<(x, \mathscr{U}(\theta); \theta) \qquad (2.2)$$

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

Since the sets $F^=(\theta^*)$ and $F_*^=(\theta)$ are generally incomparable on a region of stability, the two conditions are different and so are their Lagrange multiplier functions. Note that $S(\theta^*)$, in Theorem 2.4, is not required to be a path-connected set.

3. Modified input constraint qualifications

We introduce a new kind of ICQ's referred to below as "modified input constraint qualifications" (MICQ).

3.1 DEFINITION. Consider the convex model (P, θ), with a realistic objective function at some $\theta^* \in I$, and let $S(\theta^*)$ be a region of stability at θ^*. A condition on the constraints of the model (P, θ), with the property that for every $\theta \in N(\theta^*) \cap S(\theta^*)$ where $N(\theta^*)$ is a neighbourhood of θ^*, the system

$$f^0(x, \theta) < \tilde{f}(\theta^*)$$
$$f^i(x, \theta) < 0, \qquad i \in \mathscr{P}^<(\theta^*) \qquad\qquad (MC, \theta)$$
$$x \in F^=(\theta)$$

is inconsistent, is called a modified input constraint qualification.

Since the sets $F^=(\theta^*)$ and $F^=(\theta)$ are generally incomparable on regions of stability, ICQ's and MICQ's are essentially different conditions on the constraints of the model (P, θ). In other words, there are situations where an ICQ is not an MICQ and vice versa.

The following three conditions are easily verified to be MICQ's. There is a neighbourhood $N(\theta^*)$ of θ^* such that for every $\theta \in N(\theta^*) \cap S(\theta^*)$:

MICQ1: "$f^i(x, \theta) \leq 0, i \in \mathscr{P}^=(\theta^*) \setminus \mathscr{P}^=(\theta)$ for every $x \in F^=(\theta)$";

MICQ2: "$\mathscr{P}^=(\theta) = \mathscr{P}^=(\theta^*)$";

MICQ3: "The constraints $f^i, i \in \mathscr{P}^=(\theta^*)$ do not depend on θ".

Because of the continuity of the constraints it is obvious that MICQ3 is also MICQ2. Clearly MICQ2 is stronger than MICQ1. So it suffices to show that MICQ1 is MICQ.

3.2 LEMMA. *Consider the convex model (P, θ) with a realistic objective function at some $\theta^* \in I$. Suppose that θ^* is a locally optimal input with respect to a region of stability $S(\theta^*)$. Then there is a neighbourhood $N(\theta^*)$ of θ^* where the condition MICQ1 is a modified input constraint qualification.*

PROOF. If such a neighbourhood did not exist, we could find sequences $\theta^k \in S(\theta^*)$, $\theta^k \to \theta^*$ and $x^k = x^k(\theta^k) \in F^=(\theta^k)$ such that

$$f^0(x^k, \theta^k) < \tilde{f}(\theta^*) \qquad\qquad (3.1)$$

and

$$f^i(x^k, \theta^k) < 0, \qquad i \in \mathscr{P}^<(\theta^*).$$

But

$$f^i(x^k, \theta^k) \leq 0, \qquad i \in \mathscr{P}^=(\theta^*) \setminus \mathscr{P}^=(\theta^k)$$

by the condition MICQ1. Hence $x^k \in F(\theta^k)$ and (3.1) implies

$$\tilde{f}(\theta^k) \leq f^0(x^k, \theta^k) < \tilde{f}(\theta^*) \qquad\qquad (3.2)$$

contradicting local optimality of θ^*.

It is obvious that on any region of stability $S(\theta^*)$, where $F^=(\theta^*) \subset F^=(\theta)$, any MICQ is also ICQ. On the other hand, if on a region of stability $F^=(\theta) \subset F^=(\theta^*)$

then the opposite is true. The two conditions coincide on the regions of stability where locally $F^=(\theta) = F^=(\theta^*)$. Note that ICQ1 is not necessarily ICQ on an *arbitrary* region of stability. However, on such a region an MICQ may be satisfied.

Also note that MICQ1 always holds on the set

$$R_2(\theta^*) = \{\theta: f^i(x, \theta) \leq 0 \forall x \in F^=(\theta), \qquad i \in \mathscr{P}^=(\theta^*) \setminus \mathscr{P}^=(\theta)\} \cap I,$$

which is a region of stability, provided that the point-to-set mapping $\gamma: \theta \rightarrow F^=(\theta)$ is lower semicontinuous at θ^* (see [18]). (The region R_2 was used recently to prove stability of the Charnes-Cooper tests for efficiency of multiple-input multiple-output decision making units in data envelopment analysis, see [5].)

Now we are ready to use an MICQ to state a new necessary condition for an optimal input.

3.3 THEOREM. *Consider the convex model (P, θ) with a realistic objective function at some $\theta^* \in I$. Let $\tilde{x}(\theta^*)$ be a corresponding optimal solution and let $S(\theta^*)$ be an arbitrary path-connected region of stability at θ^*. If θ^* is a locally optimal input relative to $S(\theta^*)$, and if a modified input constraint qualification holds at θ^* relative to $S(\theta^*)$, then there exists a neighbourhood $N(\theta^*)$ of θ^* and a non-negative vector function*

$$U: N(\theta^*) \cap S(\theta^*) \rightarrow R_+^{q(\theta^*)}$$

such that, whenever $\theta \in N(\theta^*) \cap S(\theta^*)$,

$$L_*^<(\tilde{x}(\theta^*), u; \theta^*) \leq L_*^<(\tilde{x}(\theta^*), U(\theta^*); \theta^*) \leq L_*^<(x, U(\theta); \theta) \qquad (3.3)$$

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

PROOF. The proof is similar to those of Theorems 2.2 and 2.3. It has two parts: First, assuming (without loss of generality) that $\mathscr{P}^<(\theta^*) = \{1, \ldots, q(\theta^*)\}$, for an arbitrary but fixed $\theta \in N(\theta^*) \cap S(\theta^*)$, we construct the following two sets in $R^{q(\theta^*)+1}$:

$$K_1(\theta) = \left\{ y: y \geq \begin{bmatrix} f^0(x, \theta) \\ f^1(x, \theta) \\ \cdots\cdots\cdots \\ f^{q(\theta^*)}(x, \theta) \end{bmatrix} \text{ for at least one } x \in F^=(\theta) \right\}$$

and

$$K_2 = \left\{ y: y < \begin{bmatrix} \tilde{f}(\theta^*) \\ 0 \\ \cdots \\ 0 \end{bmatrix} \right\}.$$

The two sets are convex (note that $F^=(\theta)$ is convex) and, because an MICQ holds, it follows that

$$K_1(\theta) \cap K_2 = \varnothing$$

for every θ in a sufficiently small neighbourhood $N(\theta^*)$.

In the second part of the proof we separate $K_1(\theta)$ and K_2. There exists a nonzero vector $a = a(\theta) \in R^{q(\theta^*)+1}$ and some scalar $\alpha = \alpha(\theta)$ such that (we use the inner product notation)

$$(a, y^1) \geq \alpha \geq (a, y^2) \tag{3.4}$$

for all $y^1 \in K_1(\theta)$ and all $y^2 \in \operatorname{cl} K_2$ (closure of K_2). In particular, after specification of y^1 and y^2, (3.4) yields

$$a_0 \tilde{f}(\theta^*) \leq a_0 f^0(x, \theta) + \sum_{i \in \mathscr{P}^<(\theta^*)} a_i f^i(x, \theta). \tag{3.5}$$

Now we claim that $a_0 > 0$. Otherwise $a_0 = 0$ and

$$\sum_{i \in \mathscr{P}^<(\theta^*)} a_i f^i(x, \theta) \geq 0 \tag{3.6}$$

for every $x \in F^=(\theta)$. In particular, (3.6) must hold for every

$$\hat{x} \in \operatorname{rel\ int} F(\theta).$$

But for these points

$$f^i(\hat{x}, \theta) < 0, \qquad i \in \mathscr{P}^<(\theta). \tag{3.7}$$

Now we use the fact that for every small perturbation in a path-connected region of stability we must have $\mathscr{P}^<(\theta^*) \subset \mathscr{P}^<(\theta)$; see [18] and [21]. So (3.7) implies

$$f^i(\hat{x}, \theta) < 0, \qquad i \in \mathscr{P}^<(\theta^*)$$

which, together with the fact that not all a_i's are zero, contradicts (3.6). The claim is proved.

Now, after division by a_0 and using the notation

$$U = (u_i) \in R^{q(\theta^*)}, \qquad u_i = a_i/a_0, \qquad i \in \mathscr{P}^<(\theta^*),$$

we can rewrite (3.5) as

$$\tilde{f}(\theta^*) \leq L_*^<(x, U(\theta); \theta).$$

The rest of the proof (i.e., establishing the fact that $\tilde{f}(\theta^*) = L_*^<(\tilde{x}(\theta^*), U(\theta^*); \theta^*)$ and the left-hand side inequality in (3.3)), is standard and it is omitted here. (See [22] and [25] for details.)

In view of Theorem 2.3 in [22], the existence of a saddle point $(\tilde{x}(\theta^*), U(\theta^*))$, satisfying the inequalities (3.3) for every $u \in R_+^{q(\theta^*)}$ and every $x \in F^=(\theta)$, is a sufficient condition for an optimal input. This means that, in the presence of an MICQ, Theorem 3.3 gives a characterisation of an optimal input.

Note that, for a fixed $\theta = \theta^*$, the system (MC, θ) is inconsistent. In this very special case, the above characterisation (as well as the previous two) recovers a complete characterisation of an optimal solution $\tilde{x}(\theta^*)$ for the convex program (P, θ^*) from the BBZ theory.

Finally, let us note that the set of Lagrange multipliers $\Lambda(\theta)$ and $U(\theta)$ are comparable for $\theta \in N(\theta^*) \cap S(\theta^*)$, where $N(\theta^*)$ is some neighbourhood of θ^* and $S(\theta^*)$ is a path-connected region of stability. Since on such regions $\mathscr{P}^=(\theta) \subset \mathscr{P}^=(\theta^*)$, it follows that $F_*^=(\theta) \subset F^=(\theta)$ and hence $U(\theta) \subset \Lambda(\theta)$.

4. Upper semicontinuity of the Lagrange multipliers

In this section we establish conditions for upper semicontinuity of the three kinds of Lagrange multiplier functions appearing in the three necessary conditions for an optimal input, namely Λ, $\mathscr{U}$, and U. (Recall Theorems 2.3, 2.4, and 3.3.) We shall consider these multipliers as point-to-set mappings, that is

$$\Lambda: \theta \to \Lambda(\theta) = \{\lambda_i(\theta): i \in \mathscr{P}^<(\theta^*)\}$$

etc. Following Berge [4] we recall that a point-to-set mapping $\Lambda: Z \to X$, between two topological vector spaces Z and X, is upper semicontinuous at $\theta^* \in Z$ if for each open set $\mathscr{A} \subset X$, satisfying $\Lambda(\theta^*) \subset \mathscr{A}$, there exists a neighbourhood $N(\theta^*)$ of θ^* such that

$$\Lambda(\theta) \subset \mathscr{A} \quad \text{for each } \theta \in N(\theta^*).$$

The upper semicontinuity of our three mappings will follow from the two facts: (i) uniform boundedness of $\Lambda(\theta)$ for all θ's sufficiently close to θ^* and (ii) the fact that Λ is a closed mapping at θ^* (that is, for every sequence $\theta^k \to \theta^*$ and every $\lambda^k \in \Lambda(\theta^k)$, such that λ^k converges to some λ^*, it follows that $\lambda^* \in \Lambda(\theta^*)$), by the results of Hogan [10].

First we recall the result about the Lagrange multipliers Λ (appearing in Theorem 2.2) for the sake of completeness.

4.1 THEOREM [19]. *Consider the convex model (P, θ) with a realistic objective function at a locally optimal input $\theta^* \in I$ with respect to a path-connected region of stability $S(\theta^*)$.*

(i) *Then, for an arbitrary sequence $\theta^k \in S(\theta^*)$, $\theta^k \to \theta^*$, the sequence $\Lambda(\theta^k) = \{\lambda_i(\theta^k): i \in \mathscr{P}^<(\theta^*)\}$ is uniformly bounded for all sufficiently large k's.*

(ii) *If the point-to-set mapping $\gamma_*: \theta \to F_*^=(\theta)$ is lower semicontinuous at θ^*, relative to $S(\theta^*)$, then for an arbitrary sequence $\theta^k \in S(\theta^*)$, $\theta^k \to \theta^*$, the set of limit points of the sequence $\Lambda(\theta^k)$ is nonempty and it is contained in $\Lambda(\theta^*)$.*

REMARK. The first statement of the theorem also holds for the other two Lagrange multiplier functions $\mathscr{U}$ and U. The proof is essentially the same as the one used for Λ and it is based on the one recently given in [17] for particular regions of stability. The second claim remains valid also for $\mathscr{U}$ and U, but with minor modifications in the assumptions. In the case of $\mathscr{U}$ the assumption on lower semicontinuity of the mapping γ_* can be omitted, while this requirement is replaced by the one on lower semicontinuity of the mapping $\gamma: \theta \to F^=(\theta)$ in the case of U. Let us formalise these claims.

4.2 THEOREM. *Consider the convex model (P, θ) with a realistic objective function at a locally optimal input $\theta^* \in I$ with respect to a region of stability $S(\theta^*)$. Suppose that an input constraint qualification holds at θ^* relative to $S(\theta^*)$. Then, for an arbitrary sequence $\theta^k \in S(\theta^*)$, $\theta^k \to \theta^*$, the set of limit points of the sequence $\mathscr{U}(\theta^k)$ is nonempty and it is contained in $\mathscr{U}(\theta^*)$.*

PROOF. The existence of limit points follows from the statement (i) in Theorem 4.1 and the above Remark. So, suppose that for some $u^k \in \mathscr{U}(\theta^k)$ we have $u^k \to \overline{u}$, as $\theta^k \to \theta^*$. Clearly $\overline{u} \in R_+^{q(\theta^*)}$, so we can specify $u = \overline{u}$ in (2.2) to obtain

$$L_*^<(\tilde{x}(\theta^*), \overline{u}; \theta^*) \leq L_*^<(x, \mathscr{U}(\theta); \theta) \tag{4.1}$$

for every $x \in F^=(\theta^*)$, whenever $\theta \in N(\theta^*) \cap S(\theta^*)$. We will use this inequality later in the proof.

The inequalities (2.2) also give

$$L_*^<(\tilde{x}(\theta^*), u; \theta^*) \leq L_*^<(\tilde{x}(\theta^*), u^k; \theta^k)$$

after specification $x = \tilde{x}(\theta^*) \in F(\theta^*) \subset F^=(\theta^*)$, and in the limit

$$L_*^<(\tilde{x}(\theta^*), u; \theta^*) \leq L_*^<(\tilde{x}(\theta^*), \overline{u}; \theta^*) \tag{4.2}$$

for every $u \in R_+^{q(\theta^*)}$. It remains to prove that

$$L_*^<(\tilde{x}(\theta^*), \overline{u}; \theta^*) \leq L_*^<(x, \overline{u}; \theta^*) \tag{4.3}$$

for every $x \in F^=(\theta^*)$. If this was not true, we would have

$$L_*^<(\tilde{x}(\theta^*), \overline{u}; \theta^*) > L_*^<(\hat{x}, \overline{u}; \theta^*) \tag{4.4}$$

for some $\hat{x} \in F^=(\theta^*)$. This would further imply that

$$L_*^<(\tilde{x}(\theta^*), \overline{u}; \theta^*) > L_*^<(\hat{x}, u^k; \theta^k)$$

for all sufficiently large k's in some sequences $u^k \to \overline{u}$ and $\theta^k \to \theta^*$. But this contradicts (4.1). The inequalities (4.2) and (4.3) show that $(\tilde{x}(\theta^*), \overline{u})$ is a saddle point for $\theta = \theta^*$, which completes the proof.

4.3 THEOREM. *Consider the convex model* (P, θ) *with a realistic objective function at a locally optimal input* $\theta^* \in I$ *with respect to a path-connected region of stability* $S(\theta^*)$. *Suppose that a modified input constraint qualification holds at* θ^* *relative to* $S(\theta^*)$. *Also, suppose that the point-to-set mapping* $\gamma \colon \theta \to F^=(\theta)$ *is lower semicontinuous at* θ^* *relative to* $S(\theta^*)$. *Then, for an arbitrary sequence* $\theta^k \in S(\theta^*)$, $\theta^k \to \theta^*$, *the set of limit points of the sequence* $U(\theta^k)$ *is nonempty and it is contained in* $U(\theta^*)$.

PROOF. The existence of limit points follows by the same arguments as in the proof of the preceding theorem. Also, using $u^k \in U(\theta^k)$ such that $u^k \to \overline{u}$, as $\theta^k \to \theta^*$, we find that the modified (4.1) holds:

$$L_*^<(\tilde{x}(\theta^*), \overline{u}; \theta^*) \le L_*^<(x, U(\theta); \theta) \qquad (4.5)$$

for every $x \in F^=(\theta)$, whenever $\theta \in N(\theta^*) \cap S(\theta^*)$. But the inequalities (3.3) also give

$$L_*^<(\tilde{x}(\theta^*), u; \theta^*) \le L_*^<(x^k, U(\theta^k); \theta^k)$$

for every $x^k \in F(\theta^k) \subset F^=(\theta^k)$ and every $u \in R_+^{q(\theta^*)}$. For those $x^k \to \tilde{x}(\theta^*)$, as $\theta^k \to \theta^*$, this gives in the limit (4.2), which is one of the two required saddle-point inequalities.

It is left to prove that also (4.3) holds for every $x \in F^=(\theta^*)$. If this was not true, there would exist an $\hat{x} \in F^=(\theta^*)$ such that (4.4) holds. But, since the mapping γ is assumed lower semicontinuous (and hence open), we know that the sequence $\theta^k \to \theta^*$, $\theta^k \in S(\theta^*)$, for which $u^k \to \overline{u}$, generates some sequence $x^k = x^k(\theta^k) \to \hat{x}$, $x^k \in F^=(\theta^k)$. Hence it follows that the strict inequality (4.4) implies

$$L_*^<(\tilde{x}(\theta^*), \overline{u}; \theta^*) > L_*^<(x^k, u^k; \theta^k)$$

for all k's sufficiently large. But this contradicts (4.5).

5. Necessary conditions for differentiable functions

A major difficulty in obtaining necessary conditions for the model (P, θ) with differentiable functions, from the saddle-point conditions, is that the Lagrange multiplier functions are not generally continuous. This problem has been partly resolved in the preceding section and we are now in a position to move towards differentiable versions of the optimality conditions. The simplest result will again be given for the models satisfying an ICQ. In this case no extra conditions will be required. If no ICQ of any kind is required, or if an MICQ is required, the results will be proved only on particular subsets of an arbitrary region of stability $S(\theta^*)$, such as

$$S_1(\theta^*) = \{\theta \colon F(\theta^*) \subset F_*^=(\theta)\} \cap S(\theta^*)$$

and

$$S_2(\theta^*) = \{\theta : F(\theta^*) \subset F^=(\theta)\} \cap S(\theta^*).$$

For other versions of optimality conditions for the bi-convex models, proven under different assumptions (such as the "index condition") the reader is referred to [25] and [27].

We need more notation. The conditions will be stated in terms of the polar sets of points related to the unit ball, namely

$$B(\theta^*) = \left\{ \frac{\theta - \theta^*}{\|\theta - \theta^*\|} : \theta \in S(\theta^*), \theta \neq \theta^* \right\}$$

$$B_1(\theta^*) = \left\{ \frac{\theta - \theta^*}{\|\theta - \theta^*\|} : \theta \in S_1(\theta^*), \theta \neq \theta^* \right\}$$

and

$$B_2(\theta^*) = \left\{ \frac{\theta - \theta^*}{\|\theta - \theta^*\|} : \theta \in S_2(\theta^*), \theta \neq \theta^* \right\}.$$

We recall that the *polar set* of a set K is defined by

$$K^+ = \{u : (u, x) \geq 0 \text{ for every } x \in K\}$$

where $(u, x) = u^T x$ is the Euclidean inner product. It is well known that the polar is a closed convex cone regardless of the nature of K. In fact, we will use the *derived set* of $B(\theta^*)$, that is the set of all limit points of $B(\theta^*)$, as $\theta \in S(\theta^*)$, $\theta \neq \theta^*$, $\theta \to \theta^*$. We will denote this set by $B'(\theta^*)$. Also the derived sets of $B_1(\theta^*)$ and $B_2(\theta^*)$ will be used (with sequences taken from $S_1(\theta^*)$ and $S_2(\theta^*)$, respectively).

First, for the sake of completeness, we recall a result (proven in [19]) for the models that do not require any ICQ or MICQ.

5.1 THEOREM. *Consider the convex model (P, θ) at some $\theta^* \in I$ with a realistic objective function. Let $\{\tilde{x}(\theta^*), \Lambda(\theta^*)\}$ be a unique corresponding saddle point and let the functions $f^j(\tilde{x}(\theta^*), \cdot)$, $j \in \{0\} \cup \mathscr{P}^<(\theta^*)$ be differentiable. Let $S_1(\theta^*)$ be an arbitrary path-connected region of stability at θ^* and let the point-to-set mapping $\gamma_* : \theta \to F_*^=(\theta)$ be lower semicontinuous at θ^* relative to $S_1(\theta^*)$. If θ^* is a locally optimal input with respect to $S_1(\theta^*)$, then*

$$\nabla_\theta L_*^<(\tilde{x}(\theta^*), \Lambda(\theta^*); \theta)|_{\theta=\theta^*} \in \{B_1'(\theta^*)\}^+.$$

Essentially the same formula holds for the other two cases with the only change that $\Lambda(\theta^*)$ and $B_1(\theta^*)$ are substituted by the appropriate terms. However, the assumptions also change.

5.2 THEOREM. *Consider the convex model* (P, θ) *at some* $\theta^* \in I$ *with a realistic objective function. Suppose that* θ^* *is a locally optimal input with respect to a path-connected region of stability* $S_2(\theta^*)$ *and suppose that a modified input constraint qualification holds at* θ^* *relative to* $S_2(\theta^*)$. *Let* $\{\tilde{x}(\theta^*), U(\theta^*)\}$ *be a unique corresponding saddle point and let the functions* $f^j(\tilde{x}(\theta^*), \cdot)$, $j \in \{0\} \cup \mathscr{P}^<(\theta^*)$ *be differentiable. Also assume that the point-to-set mapping* $\gamma \colon \theta \to F^=(\theta)$ *is lower semicontinuous at* θ^* *relative to* $S_2(\theta^*)$. *Then*

$$\nabla_\theta L_*^<(\tilde{x}(\theta^*), U(\theta^*); \theta)|_{\theta=\theta^*} \in \{B_2'(\theta^*)\}^+.$$

PROOF. The proof is similar to the one given for Theorem 5.1. The basic difference is that it uses Theorem 4.3 (rather than Theorem 4.1) and refers to γ (instead of γ_*).

It is interesting to observe that the simplest necessary condition is obtained in the presence of an ICQ and that no assumptions on γ or γ_* are needed.

5.3 THEOREM. *Consider the convex model* (P, θ) *at some* $\theta^* \in I$ *with a realistic objective function. Suppose that* θ^* *is a locally optimal input with respect to a region of stability* $S(\theta^*)$ *and suppose that an input constraint qualification holds at* θ^* *relative to* $S(\theta^*)$. *Let* $\{\tilde{x}(\theta^*), \mathscr{U}(\theta^*)\}$ *be a unique corresponding saddle point and let the functions* $f^j(\tilde{x}(\theta^*), \cdot)$, $j \in \{0\} \cup \mathscr{P}^<(\theta^*)$ *be differentiable. Then*

$$\nabla_\theta L_*^<(\tilde{x}(\theta^*), \mathscr{U}(\theta^*); \theta)|_{\theta=\theta^*} \in \{B'(\theta^*)\}^+.$$

PROOF. Specify $x = \tilde{x}(\theta^*)$ in (2.2) to obtain

$$\tilde{f}(\theta^*) \le f^0(\tilde{x}(\theta^*), \theta) + \sum_{i \in \mathscr{P}^<(\theta^*)} \mathscr{U}_i(\theta) f^i(\tilde{x}(\theta^*), \theta) \tag{5.1}$$

for every $\theta \in N(\theta^*) \cap S(\theta^*)$, where $N(\theta^*)$ is some neighbourhood of θ^*. After adding the terms

$$-\sum_{i \in \mathscr{P}^<(\theta^*)} \mathscr{U}_i(\theta) f^i(\tilde{x}(\theta^*); \theta^*) \ge 0$$

to (5.1) and rearranging, we have

$$f^0(\tilde{x}(\theta^*), \theta) - f^0(\tilde{x}(\theta^*), \theta^*) + \sum_{i \in \mathscr{P}^<(\theta^*)} \mathscr{U}_i(\theta)[f^i(\tilde{x}(\theta^*), \theta) - f^i(\tilde{x}(\theta^*), \theta^*)] \ge 0.$$

$$\tag{5.2}$$

Let l be an arbitrary point from $B'(\theta^*)$. Then

$$l = \lim_{k \to \infty} \frac{\theta^k - \theta^*}{\|\theta^k - \theta^*\|} \tag{5.3}$$

for some sequence $\theta^k \in S(\theta^*)$, $\theta^k \to \theta^*$. But we know that for this sequence (or some subsequence) the limits

$$\lim_{k \to \infty} \mathscr{U}_i(\theta^k), \qquad i \in \mathscr{P}^<(\theta^*) \tag{5.4}$$

exist, by Theorem 4.2. It remains to use the differentiability assumption in (5.2) to write

$$\nabla f^0(\tilde{x}(\theta^*), \theta^*)(\theta - \theta^*) + \sum_{i \in \mathscr{P}^<(\theta^*)} \mathscr{U}_i(\theta) \nabla f^i(\tilde{x}(\theta^*), \theta^*)(\theta - \theta^*) + \mathscr{O}(\|\theta - \theta^*\|^2) \geq 0.$$

After division by $\|\theta - \theta^*\| > 0$, setting $\theta = \theta^k$ and letting $k \to \infty$, using (5.3) and (5.4), uniqueness of the saddle point, and the definition of the polar set, we obtain the desired formula.

6. An example

Let us illustrate some of the above ideas on a classical program of Kuhn and Tucker [13].

6.1 EXAMPLE. Consider the non-convex program

$$\text{Min } f^0 = -x_1$$

s. t.

$$f^1 = (x_1 - 1)^3 + x_2 \leq 0$$
$$f^2 = -x_1 \leq 0$$
$$f^3 = -x_2 \leq 0.$$

The optimal solution is $x_1^* = 1$, $x_2^* = 0$. A pecularity with this program is that neither a CQ nor the KKT conditions hold at the optimal solution. In particular, the classical Lagrange multipliers do not exist.

However, after the specification $x_1 = \theta$, $x_2 = x$, the above program can be considered as the convex model

$$\text{Min}_{(x)} f^0 = -\theta$$

s. t.

$$f^1 = (\theta - 1)^3 + x \leq 0$$
$$f^2 = -\theta \leq 0$$
$$f^3 = -x \leq 0.$$

The region of stability at $\theta^* = 1$ is $S(\theta^*) = [0, 1]$. We want to check whether $\theta^* = 1$ is an optimal input relative to $S(\theta^*)$, using the restricted saddle points.

First, we find that $\mathscr{P}^<(\theta^*) = \{2\}$ and hence

$$L_*^<(x, \Lambda; \theta) = -\theta[1 + \lambda_2(\theta)].$$

The point-to-set mapping $\gamma_* : \theta \to F_*^=(\theta)$ is here determined by

$$F_*^=(\theta) = \{x : 0 \leq x \leq (1 - \theta)^3\}.$$

The saddle-point characterisation from Theorem 2.2 reduces to

$$-1 \le -\theta[1 + \lambda_2(\theta)]$$

for every $0 \le \theta \le 1$. This determines the Lagrange multiplier function

$$0 \le \lambda_2(\theta) \le (1 - \theta)/\theta \qquad (6.1)$$

for $\theta \le 1$ and close to $\theta^* = 1$. In particular, one can specify

$$\lambda_2(\theta) = (1 - \theta)/\theta.$$

It is easy to verify that the Lagrange multiplier functions, satisfying the saddle-point inequality of Theorem 2.2, do not exist at any other θ, so $\theta^* = 1$ is a unique optimal input. Since both ICQ and MICQ are satisfied, the same conclusion also follows using the results such as Theorems 2.3 and 3.3. Finally, using Theorem 4.1 and the remark that follows it, we conclude that the Lagrange multipliers $\lambda_2 = \lambda_2(\theta^k)$ are uniformly bounded, as $\theta^k \in S(\theta^*)$, $\theta^k \to \theta^*$, and the limit points are contained in $\{\lambda_2(\theta^*)\}$. Indeed, as $\theta^k < 1$, $\theta^k \to 1$, all sequences chosen from the interval (6.1) are uniformly bounded sufficiently close to $\theta^* = 1$ and here they have only one common limit point that in fact equals $\lambda_2(\theta^*) = 0$.

Finally, let us note that in this situation one can also use the results for differentiable functions from Section 5.

The results of this paper are applicable to every nonlinear program (P) which, after an appropriate splitting of the variable z, can be identified as a convex model. Using Theorems 2.3, 2.4, 3.3, 5.1, 5.2 and 5.3, necessary conditions for a locally optimal solution can be formulated as necessary conditions for a locally optimal input, of the corresponding convex models, with or without an input constraint qualification.

References

[1] B. Bank, J. Guddat, D. Klatte, B. Kummer and K. Tammer, *Nonlinear Parametric Optimization*, (Academie-Verlag, Berlin, 1982).

[2] A. Ben-Israel, and B. Mond, "First-order optimality conditions for generalized convex functions: A feasible directions approach", *Utilitas Math.* **25** (1984) 249–262.

[3] A. Ben-Israel, A. Ben-Tal, and S. Zlobec, *Optimality in Nonlinear Programming: A Feasible Directions Approach*, (Wiley-Interscience, New York, 1981).

[4] C. Berge, *Espace Topologiques, Fonctions Multivogues*, (Dunod, Paris, 1959).

[5] A. Charnes, and S. Zlobec, "Stability of efficiency evaluations in data envelopment analysis", *Research Report CCS 560, Center for Cybernetic Studies, The University of Texas, Austin* (March 1987).

[6] B. Craven, "Alternative theorems", ZAMM **62** (12) (1982) 699–705.

[7] B. Craven, "Nondifferentiable optimization by smooth approximations", *Optimization* **17** (1986) 3–17.

[8] B. Craven, and V. Jeyakumar, "Alternative and duality theorems with weakened convexity", *Utilitas Math* (to appear).

[9] M. Guignard, "Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space", *SIAM J. Control Optim.* **7** (1969) 232–241.

[10] W. W. Hogan, "Point-to-set maps in mathematical programming", *SIAM Rev.* **15** (1973) 591–603.

[11] W. Karush, "Minima of Functions of Several Variables with Inequalities as Side Constraints", M. Sc. Thesis, University of Chicago, 1939.

[12] D. Klatte, "On the lower semicontinuity of optimal sets in convex parametric optimization", *Math. Programming Stud.* **10** (1979) 104–109.

[13] H. W. Kuhn, and A. W. Tucker, "Nonlinear programming", in (ed. J. Neymann) *Proc. 2nd Berkeley Symp. Math. Stat. Prob.* (1951) 481–492.

[14] O. Mangasarian, *Nonlinear Programming.* (McGraw-Hill, New York, 1969).

[15] J. P. Penot, "A new constraint qualification conditions", *J. Optim. Theory Appl.* **48** (1986) 459–468.

[16] D. W. Peterson, "A review of constraint qualifications in mathematical programming," *SIAM Rev.* **15** (1973) 639–654.

[17] J. Semple, and S. Zlobec, "Continuity of a Lagrange multiplier function in input optimization", *Math. Programming* **34** (1986) 362–369.

[18] J. Semple, and S. Zlobec, "On a necessary condition for stability in perturbed linear and convex programming." *ZOR, Series* A: *Theorie* **31** (1987) 161–172.

[19] M. van Rooyen, and S. Zlobec, "A complete characterization of optimal economic systems with respect to stable perturbations", *TWISK Rep. CSIR.* (July 1987).

[20] Q.-L. Wei, "Stability in mathematical programming in the sense of lower semicontinuity and continuity". *Journal of the Qufu Teachers College* 1 (1981). (In Chinese.)

[21] S. Zlobec, "Regions of stability for ill-posed convex programs", *Aplikace Matematiky* **27** (1982) 176–191. Addendum, Ibid **31** (1986) 109–117.

[22] S. Zlobec, "Characterizing an optimal input in perturbed convex programming," *Math. Programming* **25** (1983) 109–121; Corrigendum, *Ibid* **35** (1986) 368–371.

[23] S. Zlobec, "Input optimization: I. Optimal realizations of mathematical models", *Math. Programming* **31** (1985) 245–268.

[24] S. Zlobec, "Input optimization: III. Optimal realizations of multi-objective models", *Optimization* **17** (1986) 429–445.

[25] S. Zlobec, "Survey of input optimization", *Optimization* **18** (1987) 309–348.

[26] S. Zlobec, "Optimality conditions for a class of nonlinear programs", *Utilitas Math.* **32** (1987) 217–230.

[27] S. Zlobec, "An index condition in input optimization", *Utilitas Math.* **33** (1988) 183–192.

[28] S. Zlobec, and A. Ben-Israel, "Perturbed convex programs: Continuity of optimal solutions and optimal values", *Methods of Operations Research*, **31** (1979) 739–749.