Mathematical Programming 19 (1980) 32–60.
North-Holland Publishing Company

GEOMETRY OF OPTIMALITY CONDITIONS AND CONSTRAINT QUALIFICATIONS: THE CONVEX CASE*

Henry WOLKOWICZ

Department of Mathematics, The University of Alberta, Edmonton, Alta., Canada

Received 7 February 1979
Revised manuscript received 22 August 1979

The cones of directions of constancy are used to derive: new as well as known optimality conditions; weakest constraint qualifications; and regularization techniques, for the convex programming problem. In addition, the "badly behaved set" of constraints, i.e. the set of constraints which causes problems in the Kuhn–Tucker theory, is isolated and a computational procedure for checking whether a feasible point is regular or not is presented.

Key words; Convex Programming, Characterizations of Optimality, Constraint Qualification, Regularization, Subgradients, Faithfully Convex, Directions of Constancy.

1. Introduction

Consider the convex programming problem

$$\text{(P)}\qquad \begin{array}{l} f^0(x) \to \min \\ \text{s.t.}\quad f^k(x) \le 0, \quad k \in \mathcal{P} = \{1, 2, \dots, m\} \end{array}$$

where $f^k: X \to R$, $k \in \{0\} \cup \mathcal{P}$ are convex, not necessarily differentiable, functions and X is a locally convex linear (Hausdorff) space.

When the functions f^k are differentiable, but not necessarily convex, we get the usual Kuhn–Tucker necessary conditions for optimality:

if the feasible point x is optimal, then the system

$$\begin{cases} \nabla f^0(x) + \sum_{k \in \mathcal{P}(x)} \lambda_k \nabla f^k(x) = 0, \\ \lambda_k \ge 0 \end{cases} \tag{1.1}$$

is consistent
where $\mathcal{P}(x)$ denotes the set of binding (active) constraints at x. These conditions may fail unless a constraint qualification (regularity condition) holds at x (e.g. [20]). Many authors have given a variety of constraint qualifications (e.g. [3, 5]).

* This research was supported by the National Research Council of Canada and le Gouvernement du Quebec and is part of the author's Ph.D. Dissertation done at McGill University, Montreal, Que., under the guidance of Professor S. Zlobec.

Gould and Tolle [21] and Guignard [22] have presented a necessary and sufficient constraint qualification (weakest constraint qualification) at x, i.e. a condition on the constraints f^k, $k \in \mathcal{P}(x)$, which holds if and only if the Kuhn–Tucker conditions hold for all differentiable functions f^0 which achieve a constrained minimum at x. (Note that x is then called a regular, or Lagrange regular, point.)

The convexity assumption we have made is the natural framework for studying problem (P). We can now develop an elegant theory which does not require differentiability, produces sufficient optimality conditions and still allows many applications (see e.g. Rockafellar [27]). Furthermore, convex functions have the nice property that, if the directional derivative in the direction d is 0, then d is a direction of increase or a direction of constancy. This allows the following characterization of optimality which holds without any constraint qualification (see Ben-Israel et al. [7]):

x feasible is optimal for (P) if and only if,
for every $\Omega \subset \mathcal{P}(x)$, the system

$$\begin{cases} \nabla f^0(x) + \sum_{k \in \mathcal{P}(x) \setminus \Omega} \lambda_k \nabla f^k(x) \in \left(\bigcap_{k \in \Omega} D_k^{=}(x) \right)^* \\ \lambda_k \geq 0 \end{cases} \tag{1.2}$$

is consistent
where $D_k^{=}(x)$ is the cone of directions of constancy of f^k at x and $\cdot^*$ denotes the nonnegative polar. Following this result, Abrams and Kerzner [2] have shown that one need only consider the single set $\Omega = \mathcal{P}^=$ in (1.2) where $\mathcal{P}^=$ is the set of constraints which are identically 0 on the entire feasible set (see also [8]). Then, Ben-Tal and Ben-Israel [9] extended these results to the nondifferentiable case. In Section 5, we use the approach of Gould and Tolle [20] to derive several optimality criteria which use the cones of directions of constancy. More specifically, we show that, under certain closure conditions, $\mathcal{P}^=$ is not the only single set that can be used to characterize optimality in (1.2).

The above optimality criteria has been used to formulate algorithms that solve (P) in the absence of any constraint qualification (see e.g. [11, 31, 34]). These algorithms use the cones of directions of constancy which have been calculated in [30]. However, if x solves (P), but x is not a Kuhn–Tucker point, i.e. the Kuhn–Tucker conditions (1.1) do not hold at x, then the program (P) is "unstable", i.e. the "perturbation" function, which is the optimal value of (P) as a function of perturbations of its right-hand side, may decrease infinitely steeply in some direction. Thus, though we may solve (P), in practice our solution may be nowhere near the true solution. It is therefore of interest to know beforehand whether or not x is a Kuhn–Tucker point. Now, if a constraint qualification holds at x, then x is necessarily a Kuhn–Tucker point for all objective functions f^0 which achieve a constrained minimum at x. Program (P) is therefore "stable" at x for all such f^0. In Section 5 we present several different weakest constraint qualifications. Furthermore, in the case of faithfully convex, differentiable constraints, we show that a feasible point x is regular, i.e. some constraint qualification holds at x, if and only if every feasible point x is regular. We also show how to verify, computationally, whether or not x is regular.

In Section 2, we present several preliminaries which include showing directly that the cone of directions of constancy at x, of a continuous, faithfully convex function, is a subspace independent of $x \in X$. Section 3 introduces the set of "badly behaved constraints" at x, denoted $\mathcal{P}^b(x)$, i.e. these are the constraints which create problems in the Kuhn–Tucker theory. In fact, we will see that, the condition $\mathcal{P}^b(x) = \emptyset$ plus a certain closure condition, is a weakest constraint qualification (see Theorem 6.1). Section 4 gives several relationships between the cones of directions of constancy, the tangent cone and the linearizing cone. We conclude with several regularization techniques in Section 7. This includes regularizing program (P), when the constraints are faithfully convex so that Slater's condition holds. At the same time, this regularization reduces the number of variables and constraints of (P) (see Theorem 7.2).

2. Preliminaries

In this section we present some preliminary definitions and results needed in the sequel.

We consider the convex programming problem

$$\text{(P)} \qquad \begin{aligned} & f^0(x) \to \min \\ & \text{s.t.} \quad f^k(x) \le 0, \quad k \in \mathcal{P} = \{1, \dots, m\} \end{aligned}$$

where $f^k : X \to R$ are continuous convex functions, defined on the locally convex space X for all $k \in \{0\} \cup \mathcal{P}$. (Without loss of generality, we assume that none of the functions is constant.) Unless otherwise specified, we assume that the *feasible set*

$$S = \{x \in X : f^k(x) \le 0 \quad \text{for all} \quad k \in \mathcal{P}\}$$

is not empty. The set of *binding constraints*, at $x \in S$, is

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

An important subset of $\mathcal{P}$, independent of x, is the *equality set*

$$\mathcal{P}^= = \{k \in \mathcal{P} : f^k(x) = 0 \quad \text{for all} \quad x \in S\}.$$

(See e.g. Abrams and Kerzner [2].) This is the set of indices k for which the constraint f^k vanishes on the entire feasible set. We then denote

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

Note that unlike $\mathcal{P}^=$, $\mathcal{P}^<(x)$ depends on x.

Following Ben-Israel et al. [8], we define the relations

$$\text{“relation” is “}=\text{”, “}<\text{”, “}\leq\text{” or “}>\text{”},$$

by

$$D_f^{\text{“relation”}}(x) = \{d \in X\text{: there exists } \bar{\alpha} > 0 \text{ with } f(x+\alpha d) \text{ “relation” } f(x) \text{ for all } 0 < \alpha \leq \bar{\alpha}\}.$$

These are the cones of directions of constancy, descent, nonincrease and increase respectively. For simplicity of notation, we let

$$D_k^{\text{“relation”}}(x) = D_{f^k}^{\text{“relation”}}(x)$$

and

$$D_\Omega^{\text{“relation”}}(x) = \bigcap_{k\in\Omega} D_k^{\text{“relation”}}(x) \quad \text{for } \Omega \subset \mathcal{P}.$$

Remark 2.1. Following Rockafellar [28], we say that a convex function f is *faithfully convex* if: f is affine on a line segment only if it is affine on the whole line containing that segment. For a function f in the class of continuous faithfully convex functions, the cone $D_f^=(x)$ is a subspace independent of x. If $X = R^n$, then Rockafellar has shown that f is faithfully convex if and only if it is of the form

$$f(x) = h(Ax+b) + a\cdot x + \alpha$$

where $A \in R^{m\times n}$, $b \in R^m$, $a \in R^n$, $\alpha \in R$ and the function $h: R^m \to R$ is strictly convex. It is easy to see that $D_f^=(x) = \mathcal{N}([A/a^t])$ where $\mathcal{N}(\cdot)$ denotes null space, and is a subspace independent of x.

In the following lemma we collect some properties of the above mentioned directions. We also show directly that the cone of directions of constancy of a continuous faithfully convex function on X, is a subspace independent of $x \in X$.

Lemma 2.1. *Suppose that $f: X \to R$ is convex and continuous and $x \in S$.*

Then:

(a) *$D_f^\leq(x)$ is a convex cone, $D_f^<(x)$ is a convex blunt open cone and*

$$X = D_f^<(x) \cup D_f^=(x) \cup D_f^>(x).$$

(b) conv $D_f^=(x) \subset D_f^\leq(x)$, *where* conv *denotes convex hull, and $D_{\mathcal{P}^=}^=(x)$ is convex (see* [13] *or the proof of Lemma* 4.1(b) *below).*

(c) $D_{\mathcal{P}(x)}^\leq(x) = D_{\mathcal{P}^<(x)}^\leq(x) \cap D_{\mathcal{P}^=}^=(x)$.

(d) $D_{\mathcal{P}^=}^=(x) \cap D_{\mathcal{P}^<(x)}^<(x) \neq \emptyset$. *(Note that $D_{\mathcal{P}^<(x)}^<(x)$ is open.)*

(e) *If f is both faithfully convex and continuous, then $D_f^=(x) = D_f^=$ is a closed subspace of X, independent of x.*

Proof. For (a)–(d), see e.g. [7, 8, 13].

(e) First, let us show that $D_f^=(x)$ is a subspace. Suppose that $d_1, d_2 \in D_f^=(x)$ and let $d = d_1 + d_2$. If $\alpha \in R$, then

$$\begin{aligned} f(x+\alpha d) &= f\left(\frac{1}{2}\left(x+2\alpha d_1\right)+\frac{1}{2}\left(x+2\alpha d_2\right)\right) \\ &\le \frac{1}{2}f(x+2\alpha d_1)+\frac{1}{2}f(x+2\alpha d_2) \quad \text{since } f \text{ is convex} \\ &= f(x) \quad \text{since } d_1, d_2 \in D_f^=(x) \text{ and } f \text{ is faithfully convex.} \end{aligned}$$

Therefore f is bounded above on the whole line $x + \alpha d$, $\alpha \in R$, which implies that f is constant on this line (see e.g. Rockafellar [27, p. 69]). Thus, $d \in D_f^=(x)$. This shows that $D_f^=(x)$ is closed under addition. That $D_f^=(x)$ is closed under scalar multiplication is clear, from the definition of a faithfully convex function. This shows that $D_f^=(x)$ is a subspace. That it is closed follows from the continuity of f.

We have left to show that $D_f^=(x) = D_f^=$, i.e. it is independent of x. Suppose that $x, y \in X$ and $d \in D_f^=(x)$. We will show that $d \in D_f^=(y)$.

Case (i): Suppose that $f(y) \le f(x)$. We will first show that

$$f(y+\alpha d) \le f(x) \quad \text{for all } \alpha \in R.$$

Let $\alpha \in R$ and $1 > t_k > 0$ with $t_k \to 0$ as $k \to \infty$. Consider the directions $z^k = \alpha d + t_k(x-y)$ and let $\gamma_k = 1/t_k$. Then

$$\begin{aligned} f(y) &\le f(x) \\ &= f(x+\gamma_k \alpha d) \quad \text{since } d \in D_f^=(x) \\ &= f(y+\gamma_k z^k). \end{aligned}$$

By convexity of f and since $\gamma_k > 1$, we conclude that

$$f(y+z^k) \le f(x)$$

and thus, by continuity of f, we see that

$$f(y+\alpha d) = \lim_{k\to\infty} f(y+z^k) \le f(x).$$

This shows that f is bounded on the line $y + \alpha d$, $\alpha \in R$, and therefore, f is constant on this line, i.e. $d \in D_f^=(y)$.

Case (ii): Suppose that $f(x) < f(y)$. By a similar argument to case (*i*), we see that

$$f(y+\alpha d) = \lim_{k\to\infty} f(y+\alpha d + t_k(x-y)) \le f(y)$$

for all $\alpha \in R$, i.e. $d \in D_f^=(y)$.

We have assumed that our functions are convex, but not necessarily differentiable. Nonsmooth, or nondifferentiable, functions occur quite often in convex analysis. Applications for these functions arise in approximation theory (e.g. Dem'yanov and Malozemov [16]) duality theory (e.g. Rockafellar [27]) and semi-infinite programming (e.g. Ben-Tal et al. [10]). (See also Clarke [15] and Pshenichnyi [24].) For convex functions, it is possible to develop a complete calculus without assuming differentiability (e.g. Rockafellar [27], Pshenichnyi [24] and Holmes [23]). We now recall some concepts dealing with directional derivatives and subgradients of a convex function f, defined on the locally convex space X.

The *directional derivative* of f at x, in the direction d, is defined as

$$\nabla f(x; d) = \lim_{t \downarrow 0} \frac{f(x+td)-f(x)}{t}.$$

Convex functions have the useful property that the directional derivatives exist universally (e.g. [27, Theorem 23.1]).

A vector $\phi \in X'$ is said to be a *subgradient* of a convex function f, at the point x, if

$$f(z) \geq f(x) + \phi \cdot (z-x) \quad \text{for all } z \in X.$$

(Note that X' is the topological dual of X, equipped with the ω^*-topology.) The set of all subgradients of f at x is then called the *subdifferential* of f at x and is denoted by $\partial f(x)$.

If the directional derivative of f at x is a continuous linear functional, i.e. if $\nabla f(x; \cdot) = \phi \in X'$, then

$$\phi \cdot d = \lim_{t \to 0} \frac{f(x+td)-f(x)}{t}$$

and ϕ is called the *gradient* of f at x and denoted $\nabla f(x)$. Note that in this case

$$\partial f(x) = \{\nabla f(x)\}.$$

We collect some useful properties in the following lemma. For more details and proofs, see e.g. [23, 27].

Lemma 2.2. *Suppose that $f: X \to R$ is convex. Then*

(a) $\nabla f(x; \cdot)$ *is a finite, sublinear functional on X for all $x \in X$.*

If, in addition, f is continuous at x, then:

(b) $\nabla f(x; d) = \max\{\phi \cdot d \colon \phi \in \partial f(x)\}$; *and*

(c) $\partial f(x)$ *is a non-empty, compact convex subset of X'.*

The next lemma presents some of the relations that exist between the subgradients and the directions introduced above. For the proofs see Ben-Tal and Ben-Israel [9].

Lemma 2.3. *Suppose that $f: X \to R$ is convex. Then*
(a) $D_f^<(x) = \{d \in X: \nabla f(x; d) < 0\}$.
If $\nabla f(x)$ exists, then:
(*b*) $D_f^<(x) = \{d \in X: \nabla f(x) \cdot d < 0\}$; *and*
(c) $\operatorname{conv} D_f^=(x) = D_f^=(x) \subset \{d \in X: \nabla f(x) \cdot d = 0\}$.

We now collect some useful results on polar sets. These results can be found in e.g. Girsanov [19] and Holmes [23] (see also Borwein [12]).

Recall that for $M \subset X$, the polar of M is

$$M^* = \{\phi \in X': \phi \cdot x \geq 0 \quad \text{for all } x \in M\}.$$

M^* is then a closed convex cone in X'. However, if $M \subset X'$, then we define its polar to be

$$M^* = \{x \in X: \phi \cdot x \geq 0 \quad \text{for all } \phi \in M\}.$$

M^* is now a ω-closed convex cone in X.

Lemma 2.4. *Suppose that K and L are subsets of X and C is a subset of X'. Then*:
(a) $K \subset L$ *implies* $L^* \subset K^*$.
(b) $K^* = (\overline{\operatorname{conv} K})^*$, $C^* = (\overline{\operatorname{conv} C})^*$, $K^{**} = \overline{\operatorname{cone}} K$ *and* $C^{**} = \overline{\operatorname{cone}} C$, *where cone K denotes the convex cone generated by K.*
If, in addition, K and L are closed convex cones, then
(c) $(K \cap L)^* = \overline{K^* + L^*}$ *with*

$$(K \cap L)^* = K^* + L^* \quad \textit{if } \operatorname{int}(K) \cap L \neq \emptyset.$$

We now present some well-known definitions of cones used in mathematical programming (see e.g. Gould and Tolle [20] and Abadie [1]). However, the definitions are stated here in terms of subgradients.

By $F^0(x)$, we denote the cone of all continuous convex objective functions f^0 with the property that x minimizes f^0 over S.

For every subset Ω of $\mathcal{P}(x)$, the *linearizing cone* at $x \in S$, with respect to Ω, is

$$C_\Omega(x) = \{d \in X: \phi \cdot d \leq 0 \quad \text{for all } \phi \in \partial f^k(x) \text{ and all } k \in \Omega\}.$$

By Lemma 2.2(b), we see that

$$C_\Omega(x) = \{d \in X: \nabla f^k(x; d) \leq 0 \quad \text{for all } k \in \Omega\}.$$

The *cone of subgradients* at x is

$$B_\Omega(x) = \{\phi \in X': \phi = \sum_{k \in \Omega} \lambda_k \phi^k \text{ for some } \lambda_k \geq 0 \text{ and } \phi^k \in \partial f^k(x)\}.$$

This cone is convex and is also closed, when $0 \notin \operatorname{conv} \bigcup_{k \in \Omega} \partial f^k(x)$, i.e. when it is compactly generated. We now set

$$B_{\emptyset}(x) = \{0\}.$$

The linearizing cone and the cone of subgradient have the following dual property.

Lemma 2.5. *Suppose that $\Omega \subset \mathcal{P}$. Then*

$$\overline{B_\Omega(x)} = -C^*_\Omega(x).$$

Proof. Follows from the definition and Lemma 2.4 (b) and (c).

A further useful dual property obtained using polars is the following.

Lemma 2.6. (see [9]). *Suppose that $f: X \to R$ is a continuous convex function and $D^<_f(x) \neq \emptyset$ (equivalently $0 \notin \partial f(x)$). Then*

$$(D^{\leqq}_f(x))^* = -\text{cone}\ \partial f(x).$$

Gould and Tolle [21] used Farkas' lemma to prove lemma 2.5, for differentiable functions on R^n. Note that in the differentiable case, $B_\Omega(x)$ is finitely generated and thus closed.

For $x \in M$, where M is an arbitrary set in X, the *cone of tangents to M* at x is defined by

$$T(M, x) = \{d \in X\colon d = \lim \lambda_k(x^k - x) \text{ where } x^k \in M,\ \lambda_k \geq 0 \text{ and } x^k \to x\}.$$

This cone is closed and it is convex if M is. In fact, when M is convex, it is exactly the $\overline{\text{cone}(M-x)}$, the *support cone* of M at x. For further properties, see e.g. Guignard [22] and Holmes [23].

The cone of tangents is used in optimization theory to describe the geometry of the feasible set. For example, one gets the following characterization of optimality.

Theorem 2.1. (see [23, p. 30]). *$x \in S$ is optimal for* (P) *if and only if*

$$\partial f^0(x) \cap T^*(S, x) \neq \emptyset.$$

Note that this characterization is in terms of the feasible set, rather than the constraints.

3. The "badly behaved" constraints

For $x \in S$, let

$$\mathcal{P}^b(x) \triangleq \{k \in \mathcal{P}^=\colon (D^>_k(x) \cap C_{\mathcal{P}(x)}(x)) \setminus \overline{D^=_{\mathcal{P}^=}(x)} \neq \emptyset\}.$$

We call $\mathcal{P}^b(x)$ the set of "*badly behaved*" *constraints* at $x \in S$ for program (P).

The set $\mathcal{P}^b(x)$ is the set of constraints that creates problems in the Kuhn–Tucker theory. These are the constraints in $\mathcal{P}^=$, whose analytic properties (given by the directional derivatives) do not fully describe the geometry of the feasible set (given by the feasible directions). It will be shown in Section 6 that

$$\mathcal{P}^b(x) = \emptyset \quad \text{and} \quad B_{\mathcal{P}(x)}(x) \text{ is closed}$$

is a necessary and sufficient condition for the Kuhn–Tucker theory to hold at x, independent of f^0, i.e. it is a weakest constraint qualification.

Once $\mathcal{P}^=$ is found, then, for any given index $k_0 \in \mathcal{P}^=$, we see that $k_0 \in \mathcal{P}^b(x)$ if and only if the system

$$\begin{cases} \nabla f^{k_0}(x; d) = 0, \\ \nabla f^k(x; d) \leq 0 \quad \text{for all } k \in \mathcal{P}(x) \sim k_0, \\ \qquad d \notin D^=_{k_0}(x) \cup \overline{D^=_{\mathcal{P}^=}(x)} \end{cases}$$

is consistent. (Note that when $D^=_{k_0}(x)$ is closed, then $\overline{D^=_{\mathcal{P}^=}(x)} \subset D^=_{k_0}(x)$. This simplifies the above system and thus, the corresponding definition for the "badly behaved" set.)

The set $\mathcal{P}^b(x)$ is not equal to $\mathcal{P}^=$ in general. In fact, if

$$E_k(x) = D^=_k(x) \tag{3.1}$$

where

$$E_k(x) \triangleq \{d \in X: \nabla f^k(x; d) = 0\},$$

then f^k is "never badly behaved" at x, i.e. $k \notin \mathcal{P}^b(x)$ independent of the other constraints. This class of functions which are "never badly behaved" at x includes all continuous linear functionals on X. Furthermore, if $X = R^n$, $\nabla f(x) \neq 0$ and f is a strictly convex function of one variable, considered as a function on R^n (i.e. if the *restriction* of f to R^1 is strictly convex), then f is a nonlinear function which is "never badly behaved" at x. (See Ben-Israel et al. [7] for definitions and properties of functions whose restrictions are strictly convex.)

The class of functions which are "never badly behaved" at x also includes the "distance" functions defined below. We will see, in Section 7, that every program (P) can be "regularized" by the addition of one such "distance" function.

Lemma 3.1. *Suppose that X is a normed space, K is a convex cone in X, $x \in S$ and $k \in \mathcal{P}$. If, for $y \in X$,*

$$\begin{aligned} f^k(y) &= \operatorname{dist}(y - x, K) \\ &\triangleq \inf_{z \in K} \|(y - x) - z\|, \end{aligned} \tag{3.2}$$

then f^k is a convex function on X which is "never badly behaved" at x. Furthermore,

$$\nabla f^k(x;d)=\begin{cases}0 & \textit{if } d\in K,\\ \textit{positive} & \textit{otherwise.}\end{cases} \tag{3.3}$$

Proof. That f^k is convex follows from convexity of K and the properties of a norm.

Now let $d\in X$. Then

$$\begin{aligned}\nabla f^k(x;d)&=\lim_{t\downarrow 0}\frac{f^k(x+td)-f^k(x)}{t}\\ &=\lim_{t\downarrow 0}\frac{\operatorname{dist}(td,K)}{t}\\ &=\operatorname{dist}(d,K)\quad\text{since } K \text{ is a cone.}\end{aligned}$$

This yields (3.3) and further implies that (3.1) holds. Therefore f^k is "never badly behaved" at x.

Example 3.1. Consider the program (P) with the single constraint in one variable, $f^1(x)\le 0$, where

$$f^1(x)=\begin{cases}x^2 & \text{if } x\ge 0,\\ 0 & \text{otherwise.}\end{cases}$$

Then

$$\mathcal{P}^b(x)=\begin{cases}\{1\} & \text{if } x=0,\\ \emptyset & \text{otherwise.}\end{cases}$$

However, if

$$f^1(x)=\begin{cases}x^2+x & \text{if } x\ge 0,\\ 0 & \text{otherwise,}\end{cases}$$

then $\mathcal{P}^b(x)=\emptyset$ for all x, i.e. f^1 is not "badly behaved" at x, though $1\in\mathcal{P}^=$.

Example 3.2. Now consider the three functions

$$f^1(x)=\begin{cases}(x-1)^2 & \text{if } x\ge 1,\\ 0 & \text{otherwise,}\end{cases}$$

$$f^2(x)=\begin{cases}x^2 & \text{if } x\ge 0,\\ 0 & \text{otherwise}\end{cases}$$

and

$$f^3(x)=\begin{cases}x^2+x & \text{if } x\ge 0,\\ 0 & \text{otherwise.}\end{cases}$$

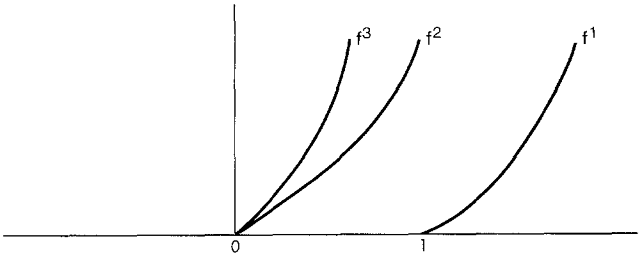

Fig. 1. Functions in Example 3.2.

If the program (P) has just the two constraints f^1 and f^2, then

$$\mathcal{P}^b(x) = \begin{cases} \{2\} & \text{if } x = 0, \\ \emptyset & \text{otherwise.} \end{cases}$$

If, however, the program (P) has all three constraints, then

$$\mathcal{P}^b(x) = \emptyset \quad \text{for all } x \in S. \tag{3.4}$$

As mentioned above, we shall see that, when $B_{\mathcal{P}(x)}(x)$ is closed, (3.4) implies that the Kuhn–Tucker theory holds, independent of the choice of the objective function f^0.

4. A basic lemma

The following lemma presents several relationships between the tangent cone, the linearizing cone and the cones of directions.

Lemma 4.1. *Suppose that $x \in S$ and the set Ω satisfies $\mathcal{P}^b(x) \subset \Omega \subset \mathcal{P}^=$. If*

$$\operatorname{conv} D^{\bar{=}}_{\Omega}(x) \text{ is closed or } \Omega = \mathcal{P}^=,$$

then:

(a) $T(S, x) = \overline{D^{\leq}_{\mathcal{P}(x)}(x)}$.

(b) $\overline{\operatorname{conv} D^{\bar{=}}_{\Omega}(x) \cap C_{\mathcal{P}(x)}(x)} = \overline{D^{\bar{=}}_{\Omega}(x)} \cap C_{\mathcal{P}(x)}(x) = \overline{D^{\bar{=}}_{\mathcal{P}^=}(x)} \cap C_{\mathcal{P}(x)}(x)$.

(c) $T(S, x) = \overline{\operatorname{conv}}\, D^{\bar{=}}_{\Omega}(x) \cap C_{\mathcal{P}(x)}(x)$.

(d) $\operatorname{conv}\{\bigcup_{k \in \mathcal{P}^<(x)} \partial f^k(x)\} \cap (D^{\bar{=}}_{\Omega}(x))^* = \emptyset$.

Proof. (Since the point $x \in S$ is fixed throughout, we will omit it in this proof when the meaning is clear. For example $C_{\mathcal{P}}$ denotes $C_{\mathcal{P}(x)}(x)$.)

(a) The result follows from the fact that

$$D^{\leq}_{\mathcal{P}(x)}(x) = \operatorname{cone}(S - x)$$

while

$$T(S, x) = \overline{\text{cone}}(S - x).$$

(b) (i) First, let us show that

$$\overline{\text{conv}}\, D^{=}_{\bar{\Omega}} \cap C_{\mathcal{P}} \subset \overline{\text{conv}\, D^{=}_{\bar{\Omega}} \cap C_{\mathcal{P}}}. \tag{4.1}$$

By hypothesis and the fact that $\overline{\text{conv}}\, D^{=}_{\mathcal{P}^{=}} \subset C_{\mathcal{P}^{=}}$ and $C_{\mathcal{P}} = C_{\mathcal{P}^{=}} \cap C_{\mathcal{P}^{<}}$, it is sufficient to show that

$$\overline{\text{conv}}\, D^{=}_{\mathcal{P}^{=}} \cap C_{\mathcal{P}^{<}} \subset \overline{\text{conv}\, D^{=}_{\mathcal{P}^{=}} \cap C_{\mathcal{P}^{<}}}.$$

But this follows since

$$\hat{d} \in D^{=}_{\mathcal{P}^{=}} \cap D^{\leq}_{\mathcal{P}^{<}} \subset \text{conv}\, D^{=}_{\mathcal{P}^{=}} \cap \text{int}\, C_{\mathcal{P}^{<}} \neq \emptyset,$$

by Lemma 2.1(d).

(ii) Next, let us show that

$$\overline{D^{=}_{\bar{\Omega}} \cap C_{\mathcal{P}}} \subset \overline{D^{=}_{\mathcal{P}^{=}}} \cap C_{\mathcal{P}}. \tag{4.2}$$

Suppose that

$$d \in (D^{=}_{\bar{\Omega}} \cap C_{\mathcal{P}}) \setminus (\overline{D^{=}_{\mathcal{P}^{=}}} \cap C_{\mathcal{P}}).$$

We will find a set $I \subset \mathcal{P}^{=}$ and feasible directions $d_\lambda \in D^{\leq}_{\mathcal{P}}$, which are directions of decrease for f^k, $k \in I$. This will contradict the definition of $\mathcal{P}^{=}$.

By the assumption, we can find a nonempty set $I \subset \mathcal{P}^{=} \setminus \Omega$ such that

$$d \in C_{\mathcal{P}} = \bigcap_{k \in \mathcal{P}} (-\partial f^k)^*, d \in D^{=}_{\mathcal{P}^{=} \setminus I}$$

but

$$d \notin D^{=}_{k} \cup \overline{D^{=}_{\mathcal{P}^{=}}} \quad \text{for each } k \in I.$$

Recall that when $k_0 \in \mathcal{P}^{=}$, then f^{k_0} is "badly behaved" at x if the system

$$\begin{cases} \nabla f^{k_0}(x; d) = 0, \\ \nabla f^{k}(x; d) \leq 0, \quad k \in \mathcal{P}(x) \setminus k_0, \\ \qquad d \notin D^{=}_{k_0} \cup \overline{D^{=}_{\mathcal{P}^{=}}} \end{cases}$$

is consistent. Therefore, since

$$I \subset \mathcal{P}^{=} \setminus \Omega \subset P^{=} \setminus \mathcal{P}^{b},$$

we see that

$$\nabla f^k(x; d) < 0 \quad \text{for all } k \in I,$$

i.e.

$$d \in D^{<}_{I} \cap D^{=}_{\mathcal{P}^{=} \setminus I}. \tag{4.3}$$

By Lemma 2.1(d), there exists

$$\hat{d} \in D^{=}_{\mathcal{P}^=} \cap D^{<}_{\mathcal{P}^<}. \tag{4.4}$$

Let

$$d_\lambda \triangleq \lambda \hat{d} + (1-\lambda) d. \tag{4.5}$$

Then, by (4.3)

$$d_\lambda \in D^{<}_{I} \quad \text{for all } 0 \leq \lambda < 1. \tag{4.6}$$

Furthermore, Lemma 2.1(b) implies that

$$d_\lambda \in D^{\leqq}_{\mathcal{P}^= \setminus I} \quad \text{for all } 0 < \lambda < 1. \tag{4.7}$$

Now, by continuity and (4.4), there exists $0 < \beta < 1$ such that

$$d_\lambda \in D^{<}_{\mathcal{P}^<} \quad \text{for all } \beta \leq \lambda < 1. \tag{4.8}$$

From (4.6), (4.7) and (4.8), we conclude that

$$d_\lambda \in D^{\leqq}_{\mathcal{P}} \cap D^{<}_{I} \quad \text{for all } \beta \leq \lambda < 1,$$

contradiction. Thus we have shown that

$$D^{=}_{\Omega} \cap C_{\mathcal{P}} \subset \overline{D^{=}_{\mathcal{P}^=}} \cap C_{\mathcal{P}}.$$

The inclusion (4.2) follows, since both $C_{\mathcal{P}}$ and $\overline{D^{=}_{\mathcal{P}^=}}$ are closed.

(iii) By a similar argument, in particular employing Lemma 2.1(b) again, we see that

$$\operatorname{conv} D^{=}_{\Omega} \cap C_{\mathcal{P}} \subset D^{=}_{\Omega} \cap C_{\mathcal{P}}. \tag{4.9}$$

(The same argument shows that $D^{=}_{\mathcal{P}^=}$ is convex, see [13].)

The desired result now follows from (4.1), (4.2), (4.9) and

$$\overline{D^{=}_{\mathcal{P}^=}} \cap C_{\mathcal{P}} \subset \overline{D^{=}_{\Omega}} \cap C_{\mathcal{P}} \subset \overline{\operatorname{conv}} \, D^{=}_{\Omega} \cap C_{\mathcal{P}}.$$

(c) By (a), (b), and (4.1) it is sufficient to show that

$$\overline{D^{\leqq}_{\mathcal{P}}} = \overline{\operatorname{conv} D^{=}_{\mathcal{P}^=} \cap C_{\mathcal{P}}}.$$

That

$$\overline{D^{\leqq}_{\mathcal{P}}} \subset \overline{\operatorname{conv} D^{=}_{\mathcal{P}^=} \cap C_{\mathcal{P}}} \tag{4.10}$$

is clear from the definitions and Lemma 2.1(c). To prove the converse, we first show that

$$\operatorname{conv} D^{=}_{\mathcal{P}^=} \cap C_{\mathcal{P}} \subset \overline{D^{\leqq}_{\mathcal{P}}}. \tag{4.11}$$

Suppose that we are given

$$d \in \operatorname{conv} D^{=}_{\mathcal{P}^=} \cap C_{\mathcal{P}}$$

and the neighbourhood of the origin, n. We need to show that

$$D^{\leq}_{\mathscr{P}} \cap (n+d) \neq \emptyset. \tag{4.12}$$

Let $\hat{d}$ and d_λ be defined as in (4.4) and (4.5) respectively. Then

$$d_\lambda \cdot \phi = \lambda \hat{d} \cdot \phi + (1-\lambda) d \cdot \phi < 0,$$

for all $0 < \lambda \leq 1$ and all $\phi \in \bigcup_{k \in \mathscr{P}^<} \partial f^k$. Therefore

$$d_\lambda \in D^{<}_{\mathscr{P}^<} \cap \operatorname{conv} D^{=}_{\mathscr{P}^=} \subset D^{\leq}_{\mathscr{P}} \quad \text{for all } 0 < \lambda \leq 1.$$

Furthermore, $d_\lambda \in n + d$ for sufficiently small λ. This proves (4.12) and thus (4.11). The desired result now follows since $\overline{D^{\leq}_{\mathscr{P}}}$ is closed.

(d) Let

$$C \triangleq \operatorname{conv}\left\{ \bigcup_{k \in \mathscr{P}^<} \partial f^k \right\}. \tag{4.13}$$

Suppose that the intersection is not empty. Then $\mathscr{P}^< \neq \emptyset$ and there exists

$$\phi \in C \cap (D^{=}_{\Omega})^*$$

where $\phi = \sum_{k \in \mathscr{P}^<} \lambda_k \phi^k$ for some $\lambda_k \geq 0$, $\sum \lambda_k = 1$ and $\phi^k \in \partial f^k$. By Lemma 2.4(a) and (b),

$$D^{=}_{\Omega} \subset \{\phi\}^* \quad \text{and} \quad -C^* \subset -\{\phi\}^*.$$

Therefore

$$D^{=}_{\Omega} \cap -C^* \subset \{\phi\}^\perp \quad \text{(the annihilator of } \phi\text{)}.$$

Let $\hat{d}$ be as in (4.4). Then

$$\hat{d} \in D^{=}_{\mathscr{P}^=} \cap D^{<}_{\mathscr{P}^<} \subset D^{=}_{\Omega} \cap -C^* \subset \{\phi\}^\perp,$$

i.e. $\hat{d} \cdot \phi = 0$. But

$$\hat{d} \cdot \phi = \hat{d} \cdot \sum_{k \in \mathscr{P}^<} \lambda_k \phi^k < 0,$$

since $\hat{d} \in D^{<}_{\mathscr{P}^<}$, $\lambda_k \geq 0$ and $\sum \lambda_k = 1$. Contradiction.

The above lemma will be used to prove the optimality criteria and the necessary and sufficient constraint qualifications in the following sections.

5. Optimality criteria

In this section we present the optimality criteria, for the convex program (P), of the type:

$x \in S$ is optimal if and only if the system

$$\begin{cases} 0 \in \partial f^0(x) + \sum_{k \in \mathcal{P}(x)} \lambda_k \partial f^k(x) - G, \\ \lambda_k \geq 0 \end{cases}$$

is consistent,
where G is some nonempty cone in X'. But, first we prove the following lemma.

Lemma 5.1. *Suppose that $x \in S$ and $G \subset X'$. Then the statement: "x is optimal for (P) if and only if the system*

$$\begin{cases} 0 \in \partial f^0(x) + \sum_{k \in \mathcal{P}(x)} \lambda_k \partial f^k(x) - G, \\ \lambda_k \geq 0 \end{cases} \tag{5.1}$$

is consistent",
holds, for any fixed objective function f^0, if and only if G satisfies

$$T^*(S, x) = -B_{\mathcal{P}(x)}(x) + G. \tag{5.2}$$

Proof. Sufficiency: Suppose that G satisfies (5.2). By Theorem 2.1, we know that x is optimal if and only if $\partial f^0(x) \cap T^*(S, x) \neq \emptyset$. By (5.2) this implies that x is optimal if and only if $\partial f^0(x) \cap (-B_{\mathcal{P}(x)}(x) + G) \neq \emptyset$, i.e. if and only if (5.1) is consistent.

Necessity: We need to show that (5.2) holds. Suppose that $\phi \in T^*(S, x)$ and f^0 is defined by the linear functional $\phi(\cdot)$ on X. Then $\phi \in \partial f^0(x) \cap T^*(S, x)$ and we can conclude that x is optimal for (P), i.e. $\phi = f^0 \in F^0(x)$. Therefore, by the conditions (5.1) we see that $\phi \in -B_{\mathcal{P}(x)}(x) + G$. Thus

$$T^*(S, x) \subset -B_{\mathcal{P}(x)}(x) + G.$$

Conversely, let $\phi \in -B_{\mathcal{P}(x)}(x) + G$. Then we can find $\lambda_k \geq 0$ and $\phi^k \in \partial f^k(x)$ such that

$$\phi + \sum_{k \in \mathcal{P}(x)} \lambda_k \phi^k \in G.$$

Again we let f^0 be the linear functional ϕ. Then $\phi = f^0 \in F^0(x)$ by (5.1). Since $\partial f^0(x) = \{\phi\}$, Theorem 2.1 implies that $\phi \in T^*(S, x)$. Thus

$$-B_{\mathcal{P}(x)}(x) + G \subset T^*(S, x).$$

When $B_{\mathcal{P}(x)}(x)$ is closed, we see that (5.2) becomes, by Lemma 2.5,

$$T^*(S, x) = C^*_{\mathcal{P}(x)}(x) + G.$$

This condition was studied by Gould and Tolle [20] in the case when $X = R^n$ and the functions f^k are differentiable but not necessarily convex. (Note that by Lemma 2.5 and Lemma 4.1(d), we get that $B_{\mathcal{P}(x)}(x)$ is closed when the constraints f^k, $k \in \mathcal{P}^=$, are differentiable.)

By specifying G in (5.2) we get necessary and sufficient conditions for optimality. One obvious candidate for G is $T^*(S, x) \setminus C^*_{\mathcal{P}(x)}(x) \cup \{0\}$. By Lemma

4.1(a), another candidate is $(D^{\leq}_{\mathcal{P}(x)}(x))^*$. More useful candidates for G are given in the next theorem.

Theorem 5.1. *Suppose that $x \in S$, the set Ω satisfies $\mathcal{P}^b(x) \subset \Omega \subset \mathcal{P}^=$ and both*

$$\operatorname{conv} D^{=}_{\Omega}(x) \quad \textit{and} \quad -B_{\mathcal{P}(x)}(x) + (D^{=}_{\Omega}(x))^* \textit{ are closed.} \tag{5.3}$$

Then, x is optimal for (P) if and only if the system

$$\begin{cases} 0 \in \partial f^0(x) + \sum_{k \in \mathcal{P}(x)} \lambda_k \partial f^k(x) - (D^{=}_{\Omega}(x))^*, \\ \lambda_k \geq 0 \end{cases} \tag{5.4}$$

is consistent.

Proof. The result follows from Lemmas 2.4(c), 2.5, 4.1(c) and 5.1.

We have assumed that the two sets $\operatorname{conv} D^{=}_{\Omega}$ and $-B_{\mathcal{P}(x)}(x) + (D^{=}_{\Omega}(x))^*$ are closed. (This can be considered as a kind of constraint qualification.) The sets are closed, for example, when the constraints are faithfully convex and differentiable. For then both cones in the sum are polyhedral. The following two examples show that the closure assumptions are necessary.

Example 5.1. Consider the program

$$f^0(x) \to min$$
$$\text{s.t.} \quad f^k(x) \leq 0, \quad k \in \mathcal{P} = \{1, 2, 3\}$$

where $x = (x_i) \in R^3$, $f^1(x) = x_1$, $f^2(x) = -x_1$, $f^3(x) = (\text{dist}\,(x, K))^2$ (see (3.2)) and K is the self-polar, "ice-cream" cone

$$K \triangleq \{x \in R^3 : x_1 + x_2 \geq 0, \quad 2x_1x_2 \geq x_3^2\}.$$

Note that now

$$\nabla f^3(0; d) = \lim_{t \downarrow 0} \frac{\inf_{z \in k} \|td - z\|^2}{t} = 0 \quad \text{for all } d \in R^3.$$

Let $\bar{x} = 0$. Then $\bar{x} \in S$, $\mathcal{P}^= = \mathcal{P}$ while $\mathcal{P}^b(\bar{x}) = \{3\}$. Furthermore,

$$-B_{\mathcal{P}(0)}(0) = C^*_{\mathcal{P}(0)}(0) = \operatorname{span}\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}\right\}$$

and $(D^{=}_{\mathcal{P}^b(0)}(0))^* = K$. Let us show that

$$C^*_{\mathcal{P}(0)}(0) + (D^{=}_{\mathcal{P}^b(0)}(0))^* \text{ is not closed.}$$

Choose

$$k^i = \begin{pmatrix} i \\ \frac{1}{i} \\ 1 \end{pmatrix} \in K \quad \text{and} \quad l^i = \begin{pmatrix} -i \\ 0 \\ 0 \end{pmatrix} \in C^*_{\mathcal{P}(0)}(0), \quad i = 1, 2, \ldots.$$

Then

$$k^i + l^i \to \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \notin C^*_{\mathcal{P}(0)}(0) + (D^{=}_{\mathcal{P}^b(0)}(0))^*.$$

Example 5.2. Consider the program

$$\begin{aligned} &f^0(x) \to \min \\ \text{s.t.} \quad &f^k(x) \le 0, \quad k \in \mathcal{P} = \{1, 2\} \end{aligned}$$

where $x = (x_i) \in R^2$,

$$f^1(x) = \begin{cases} (x_1^2 + x_2^2 - 1)^2 & \text{if } x_1^2 + x_2^2 - 1 \ge 0, \\ 0 & \text{otherwise,} \end{cases}$$

$f^2(x) = \operatorname{dist}(x - \bar{x}, K)$, $K = \{x \in R^2 : x_1 \ge 0,\ x_2 \ge 0\}$ and $\bar{x} = (1,0)^t$. Then $S = \{\bar{x}\}$, $\mathcal{P}^= = \mathcal{P}$ while $\mathcal{P}^b(\bar{x}) = \{1\}$. Let $\Omega = \{1\}$. Then

$$D^{=}_{\Omega}(\bar{x}) = \{d \in R^2 : d_1 < 0\} \cup \{0\}$$

is not closed. Furthermore

$$T^*(S, \bar{x}) = R^2 \quad \text{and} \quad C^*_{\mathcal{P}(\bar{x})}(\bar{x}) = K.$$

Therefore

$$(D^{=}_{\Omega}(\bar{x}))^* - B_{\mathcal{P}(\bar{x})}(\bar{x}) \subset (D^{=}_{\Omega}(\bar{x}))^* + C^*_{\mathcal{P}(\bar{x})}(\bar{x}) = \{x \in R^2 : x_2 \ge 0\} \subsetneq T^*(S, x).$$

This implies that (5.2) fails and that we cannot choose $\Omega = \{1\}$ in Theorem 5.1.

The case $\Omega = \mathcal{P}^=$ in the above theorem is exceptional. In this case we no longer need the hypothesis (5.1). In fact,

$$(D^{=}_{\Omega}(x))^* - B_{\mathcal{P}(x)}(x) = T^*(S, x)$$

always holds.

Theorem 5.2. *Let $x \in S$. Then x is optimal for (P) if and only if the system*

$$\begin{cases} 0 \in \partial f^0(x) + \sum_{k \in \mathcal{P}(x)} \lambda_k \partial f^k(x) - (D^{=}_{\mathcal{P}^=}(x))^*, \\ \\ \lambda_k \ge 0 \end{cases} \tag{5.5}$$

is consistent.

Proof. First note that

$$\begin{aligned} -B_{\mathscr{P}(x)}(x) &\subset C^*_{\mathscr{P}(x)}(x) \quad \text{by Lemmas 2.5 and 2.4(a)}, \\ &\subset T^*(S, x) \quad \text{by Lemmas 4.1(c) and 2.4(a)}. \end{aligned} \tag{5.6}$$

Now

$$\begin{aligned} T^*(S,x) &= (D^{\leqq}_{\mathscr{P}(x)}(x))^* \quad \text{by Lemma 4.1(a)} \\ &= (D^{\overline{=}}_{\mathscr{P}^=}(x) \cap D^{\leqq}_{\mathscr{P}^<(x)}(x))^* \quad \text{by Lemma 2.1(c)} \\ &= (D^{\overline{=}}_{\mathscr{P}^=}(x))^* \\ &\quad + \sum_{k \in \mathscr{P}^<(x)} (D^{\leqq}_k(x))^* \quad \text{by Lemmas 2.1(b), (d) and 2.4(c)} \\ &= (D^{\overline{=}}_{\mathscr{P}^=}(x))^* - B_{\mathscr{P}^<(x)}(x) \quad \text{by Lemma 2.6} \\ &= (D^{\overline{=}}_{\mathscr{P}^=}(x))^* - B_{\mathscr{P}(x)}(x) \quad \text{by (5.6).} \end{aligned}$$

The result now follows from Lemma 5.1.

The above result is equivalent to the characterization of optimality given in [2, 7, 8, 9], the difference being $\mathscr{P}^<(x)$ replaces the set $\mathscr{P}(x)$ in (5.5). It is of interest to note that Theorem 5.1 shows that, under certain closure conditions, $\mathscr{P}^=$ can be replaced in (5.5) by any set Ω which satisfies $\mathscr{P}^b(x) \subset \Omega \subset \mathscr{P}^=$. However, it appears that $\mathscr{P}^=$ is the only set that does not require any additional closure conditions. The closure conditions, however, are always satisfied in the faithfully convex, differentiable case.

6. Weakest constraint qualifications.

A point $x \in S$ is called a *Kuhn–Tucker point* for (P) if the Kuhn–Tucker conditions are satisfied at x, i.e. if the system

$$\begin{cases} 0 \in \partial f^0(x) + \sum_{k \in \mathscr{P}(x)} \lambda_k \partial f^k(x), \\ \\ \lambda_k \geq 0 \end{cases} \tag{6.1}$$

is consistent. It is well-known that if $x \in S$ is a Kuhn–Tucker point, then x solves program (P). However, the converse does not hold in general.

We call $x \in S$ a regular point (Lagrange regular point), if the Kuhn–Tucker conditions (6.1) hold for every $f^0 \in F^0(x)$, i.e. if we can choose $G = \{0\}$ in (5.2). A *constraint qualification* is then a condition on the set of constraints which guarantees that x is a regular point. A *weakest constraint qualification* (WCQ) is a constraint qualification that holds if and only if x is a regular point. In other words, it is a condition, on the constraints, which holds at x if and only if the Kuhn–Tucker conditions characterize optimality at x for any given f^0. Gould and

Tolle [20, 21] have shown that in their setting, i.e. in the differentiable case on R^n,

$$T^*(S,x) = C^*_{\mathcal{P}(x)}(x) \quad \text{is a WCQ.} \tag{6.2}$$

Since in our setting $B_{\mathcal{P}(x)}(x)$ is not closed in general, we no longer have (6.2). However, by Lemma 5.1, we do have that

$$T^*(S,x) = -B_{\mathcal{P}(x)}(x) \quad \text{is a WCQ.}$$

By Lemma 2.5, this is equivalent to,

$$B_{\mathcal{P}(x)}(x) \text{ is closed} \quad \text{and} \quad T^*(S,x) = C^*_{\mathcal{P}(x)}(x) \text{ is a WCQ.} \tag{6.3}$$

In this section we present several different WCQ's. We also show (see Corollary 6.1) how, in the differentiable faithfully convex case, one can verify computationally whether or not x is a regular point.

Theorem 6.1. *Suppose that $x \in S$. Then T.F.A.E.:*

(a) *x is a regular point.*
(b) *$T(S,x) = C_{\mathcal{P}(x)}(x)$ and $B_{\mathcal{P}(x)}(x)$ is closed.*
(c) *$\mathcal{P}^b(x) = \emptyset$ and $B_{\mathcal{P}(x)}(x)$ is closed.*
(d) *$C_{\mathcal{P}(x)}(x) \subset \overline{D^=_{\mathcal{P}^=}(x)}$ and $B_{\mathcal{P}(x)}(x)$ is closed.*

Proof. That (a) is equivalent to (b) follows directly from (6.3) and Lemma 2.4(b).

Now, suppose that $\mathcal{P}^b(x) = \emptyset$. Then, Lemma 4.1(c) implies that

$$T(S,x) = C_{\mathcal{P}(x)}(x).$$

Conversely, suppose that $\mathcal{P}^b(x) \neq \emptyset$. Recall that $k \in \mathcal{P}^b(x)$ if $k \in \mathcal{P}^=$ and there exists

$$d \in (D^>_k(x) \cap C_{\mathcal{P}(x)}(x)) \setminus \overline{D^=_{\mathcal{P}^=}(x)}.$$

But, this implies that

$$d \notin \overline{D^=_{\mathcal{P}^=}(x)} \cap C_{\mathcal{P}(x)}(x) = T(S,x) \quad \text{by Lemma 4.1(b) and (c).}$$

Therefore,

$$T(S,x) \neq C_{\mathcal{P}(x)}(x).$$

This proves (b) is equivalent to (c).

Finally, that (b) is equivalent to (d) follows from Lemma 4.1(b) and (c).

Remark 6.1. Suppose that we can find $\hat{x} \in S$ and $\Omega \subset \mathcal{P}$ such that $f^k(\hat{x}) < 0$, for all $k \in \mathcal{P} \setminus \Omega$, and f^k is "never badly behaved" at x for all $k \in \Omega$ and some $x \in S$, i.e.

$$E_k(x) = D^=_k(x) \quad \text{for all } k \in \Omega.$$

Then, since $\mathcal{P}^b(x) \subset \mathcal{P}^=$, this implies that $\mathcal{P}^b(x) = \emptyset$, i.e. if $B_{\mathcal{P}(x)}(x)$ is closed then $x \in S$ is a regular point. In particular, in the differentiable case, we see that, when checking if Slater's condition holds, we need not worry about the functions which are "never badly behaved". In particular, we can ignore all linear functionals.

Remark 6.2. Suppose that S contains two distinct points. Then Slater's condition is a WCQ with respect to the Fritz John optimality conditions.

Proof. The Fritz John optimality conditions state that: $x \in S$ is optimal if and only if the system

$$\begin{cases} 0 \in \sum_{k \in \mathcal{P}(x) \cup \{0\}} \lambda_k \partial f^k(x), \\ \lambda_k \geq 0, \quad \sum_{k \in \mathcal{P}(x) \cup \{0\}} \lambda_k = 1 \end{cases} \tag{6.4}$$

is consistent. Necessity always holds. We need to show that, if S contains two distinct points (note that when $S = \{x\}$, then x is optimal for any f^0 chosen, the Fritz John optimality conditions hold, but Slater's condition fails), then the Fritz John conditions are sufficient for optimality, independent of the objective function f^0, if and only if Slater's condition holds. So, suppose that $x \in S$ and the system (6.4) is consistent. Now, consider the system

$$\begin{cases} \sum_{k \in \mathcal{P}(x)} \lambda_k \phi^k = 0, \\ \phi^k \in \partial f^k(x),\ \lambda_k \geq 0, \quad \sum_{k \in \mathcal{P}(x)} \lambda_k = 1. \end{cases} \tag{6.5}$$

Since the Kuhn–Tucker conditions are always sufficient for optimality and since (6.5) is independent of f^0, we see that the Fritz John conditions (6.4) are sufficient for optimality if and only if the system (6.5) is inconsistent. But (6.5) is inconsistent if and only if the system

$$\begin{cases} \phi^k \cdot d < 0 & \text{for all } k \in \mathcal{P}(x), \\ \phi^k \in \partial f^k(x), & d \in X \end{cases}$$

is consistent (Motzkin's Theorem of the Alternative [33]) if and only if Slater's condition holds.

Now, suppose that $\bar{x} \in S$ and $\nabla f^k(\bar{x})$ exists for all $k \in \mathcal{P}^=$. Consider the convex program with the linear constraints

$$g^k(x) = \nabla f^k(\bar{x}) \cdot x \leq 0, \quad k \in \mathcal{P}^=. \tag{6.6}$$

Let $\mathcal{R}^=$ denote the equality set for these constraints, i.e.

$$\mathcal{R}^= = \{k_0 \in \mathcal{P}^=: g^{k_0}(x) = 0 \text{ for all } x \in \{x \in X: g^k(x) \leq 0 \text{ for all } k \in \mathcal{P}^=\}\}. \tag{6.7}$$

(Note that since the g^k are linear, we get that

$$D^{=}_{\mathscr{R}^=} = \bigcap_{k\in\mathscr{R}^=} (\nabla f^k(\bar{x})^{\perp}.)$$

Recall that when f^k is faithfully convex, then $D^{=}_k$ is a subspace independent of x. In this case, we get the following WCQ which can be verified computationally.

Corollary 6.1. *Let $\bar{x}\in S$. Suppose that the constraints f^k, $k\in\mathscr{P}(\bar{x})$, are differentiable and f^k, $k\in\mathscr{P}^=$, are faithfully convex. Then T.F.A.E.:*

(a) *$\bar{x}$ is a regular point.*

(b) *Every $x\in S$ is a regular point.*

(c) *$D^{=}_{\mathscr{R}^=} = D^{=}_{\mathscr{P}^=}$, where $\mathscr{R}^=$ is given in* (6.7).

Proof. Consider the convex program with the linear constraints

$$g^k(x) = \nabla f^k(\bar{x})\cdot x \le 0, \quad k\in\mathscr{P}(\bar{x}) \tag{6.8}$$

and suppose that k_0 is in the equality set for these constraints, i.e. $\nabla f^k(\bar{x})\cdot d\le 0$ for all $k\in\mathscr{P}(\bar{x})$ implies that $\nabla f^{k_0}(\bar{x})\cdot d = 0$. Therefore

$$D^{\le}_{\mathscr{P}(\bar{x})}(\bar{x}) \cap D^{<}_{k_0}(\bar{x}) = \emptyset.$$

This implies that $k_0\in\mathscr{P}^=$. Furthermore, since $\mathscr{P}^=\subset\mathscr{P}(\bar{x})$, we get that

$$\mathscr{R}^= \subset \mathscr{P}^=$$

and $\mathscr{R}^=$ is the equality set for the constraints (6.8) as well as for the constraints (6.6). We now conclude that

$$\begin{aligned} D^{=}_{\mathscr{R}^=} &= \operatorname{span}\{d\in X\colon \nabla f^k(\bar{x})\cdot d\le 0 \text{ for all } k\in\mathscr{P}(\bar{x})\} \\ &= \operatorname{span} C_{\mathscr{P}(\bar{x})}(\bar{x}) \end{aligned} \tag{6.9}$$

and

$$D^{=}_{\mathscr{P}^=} \subset D^{=}_{\mathscr{R}^=}. \tag{6.10}$$

But since $D^{=}_{\mathscr{P}^=}$ is a closed subspace, we get, from Theorem 6.1, that $\bar{x}$ is a regular point if and only if

$$\operatorname{span} C_{\mathscr{P}(\bar{x})}(\bar{x}) \subset D^{=}_{\mathscr{P}^=}.$$

That (a) is equivalent to (c) now follows from (6.9) and (6.10).

Now let us show that (a) is equivalent to (b). Let $x\in S$. Then, for $k\in\mathscr{P}^=$ we see that

$$\begin{aligned} \nabla f^k(x; d) &= \nabla f^k(x)\cdot d \\ &= \lim_{t\downarrow 0} \frac{f^k(x+td)-f^k(x)}{t} \\ &= \lim_{t\downarrow 0} \frac{f^k(\bar{x}+\bar{d}+td)-f^k(\bar{x}+\bar{d})}{t} \quad \text{where } \bar{d} = x-\bar{x} \end{aligned}$$

$$= \lim_{t \downarrow 0} \frac{f^k(\bar{x} + td) - f^k(\bar{x})}{t} \quad \text{since } k \in \mathcal{P}^= \text{ implies that } \bar{d} \in D^{\bar{=}}_{\mathcal{P}^=}$$

$$= \nabla f^k(\bar{x}) \cdot d.$$

This implies that the directional derivatives $\nabla f^k(x) \cdot d$ and thus also the gradients $\nabla f^k(x)$ are independent of $x \in S$ for all $k \in \mathcal{P}^=$. Thus $D^{\bar{=}}_{\mathcal{R}^=}$ is independent of $x \in S$. The result now follows since we have already shown that (a) is equivalent to (c).

The above result shows that x is a regular point if and only if the gradients of the binding constraints supply sufficient information about the feasible directions. More precisely, since f^k, $k \in \mathcal{P}$, is convex, we can always determine the directions of decrease using the gradient, with the formula

$$D^<_k(x) = \{d \in X \colon \nabla f^k(x) \cdot d < 0\}.$$

The above theorem states that we need also be able to determine the directions of constancy $D^{\bar{=}}_{\mathcal{P}^=}(x)$ using the gradients.

Example 6.1. Suppose $S \subset R^5$ is defined by the constraints

$$\begin{aligned}
f^1(x) &= e^{x_1} + x_2^2 - 1 \le 0,\\
f^2(x) &= x_1^2 + x_2^2 + e^{-x_3} - 1 \le 0,\\
f^3(x) &= x_1 + x_4^2 + x_5^2 - 1 \le 0,\\
f^4(x) &= e^{-x_2} - 1 \le 0,\\
f^5(x) &= (x_1 - 1)^2 + x_2^2 - 1 \le 0,\\
f^6(x) &= x_1 + e^{-x_4} - 1 \le 0,\\
f^7(x) &\quad x_2 + e^{-x_5} - 1 \le 0.
\end{aligned}$$

Let us check whether every $x \in S$ is regular, while finding $\mathcal{P}^=$ and $D^{\bar{=}}_{\mathcal{P}^=}$. (The algorithm for finding $\mathcal{P}^=$ was given in [2] and later modified for faithfully convex constraints in [31].)

Initialization. Let $\bar{x} = (0, 0, 1, \frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2})$ be the chosen feasible point. Then $P_0 = A_0 = I_{5\times 5}$, $\mathcal{P}_0 = \mathcal{P}(\bar{x}) = \{1, 3, 4, 5\}$ and $\mathcal{P}_0^= = \emptyset$. The corresponding gradients are

$$\begin{aligned}
\nabla f^1(\bar{x}) &= (1, 0, 0, 0, 0),\\
\nabla f^3(\bar{x}) &= (1, 0, 0, \sqrt{2}, \sqrt{2}),\\
\nabla f^4(\bar{x}) &= (0, -1, 0, 0, 0),\\
\nabla f^5(\bar{x}) &= (-2, 0, 0, 0, 0).
\end{aligned}$$

(We simultaneously find $\mathcal{R}^=$ and $D^{\bar{=}}_{\mathcal{P}^=}$, where $\mathcal{R}^=$ is the equality set for the linear constraints

$$\nabla f^k(\bar{x}) \cdot x \le 0, \quad k \in \mathcal{P}_0 = \{1, 3, 4, 5\}.) \tag{6.11}$$

Step 0: Since $\nabla f^k(\bar{x})A_0 = \nabla f^k(\bar{x}) \neq 0$ for all $k \in \mathcal{P}_0$, we solve the system

$$\lambda_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 1 \\ 0 \\ 0 \\ \sqrt{2} \\ \sqrt{2} \end{bmatrix} + \lambda_4 \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_5 \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\lambda_1 + \lambda_3 + \lambda_4 + \lambda_5 = 1 \quad \text{and} \quad \lambda_k \geq 0.$$

A solution is $\lambda_1 = \frac{2}{3}$, $\lambda_3 = \lambda_4 = 0$ and $\lambda_5 = \frac{1}{3}$. Therefore, $J_0 = \{1, 5\}$, $\mathcal{P}_1 = \{3, 4\}$, $\mathcal{P}_1^{=} = \{1, 5\}$,

$$A_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

with $\mathcal{R}(A_1) = \bigcap_{k \in J_0} D^{=}_{f^k o P_0}$ and $P_1 = P_0 A_1 = A_1$.

(For the constraints (6.11), we see that

$$A_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

with $\mathcal{R}(A_1) = \bigcap_{k \in J_0} \nabla(f^k(\bar{x}))^{\perp}$

and $P_1 = A_1$.)

Step 1: Since $\nabla f^4(\bar{x}) P_1 = 0$ while $\nabla f^3(\bar{x}) P_1 \neq 0$ we get that

$J_1 = \{4\}$, $\mathcal{P}_2 = \{3\}$ $\mathcal{P}_2^{=} = \{1, 4, 5\}$,

$$A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

with $\mathcal{R}(A_2) = D^{=}_{f^4 o P_1}$ and $P_2 = P_1 A_2 = A_1$.

(For the constraints (6.11), we see that the corresponding system is inconsistent. Thus

$$\mathcal{R}^{=} = \{1, 5\}; \quad D^{=}_{\mathcal{R}^{=}} = \mathcal{R}\left(\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right).$$

Step 2: Since $\mathcal{P}_2 = \{3\}$ and $\nabla f^3(\bar{x})P_2 \neq 0$ we stop.

Conclusion.

$$\mathcal{P}^= = \mathcal{P}_2^= = \{1, 4, 5\}$$

and

$$\mathcal{D}_{\mathcal{P}^=}^= = \mathcal{R}(P_2) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ d_3 \\ d_4 \\ d_5 \end{pmatrix} \in R^5 \colon d_3, d_4, d_5 \in R \right\}.$$

Thus $D_{\mathcal{R}^=}^= \neq D_{\mathcal{P}^=}^=$ which implies that the points $x \in S$ are not regular.

Remark 6.3. As mentioned in the introduction, constraint qualifications are important when dealing with stability. In fact [18, Theorem 1] a solution $\bar{x}$ of (P) is a Kuhn–Tucker point if and only if the optimal value $\mu = f(\bar{x})$ is stable with respect to perturbations of the right-hand sides of the constraints. More precisely, if

$$\mu(\epsilon) = \inf\{f^0(x) \colon f^k(x) \leq \epsilon_k \quad \text{for all } k \in \mathcal{P}\}$$

denotes the perturbation function, where $\epsilon = (\epsilon_k) \in R^m$ is the perturbation vector, and if there exists a Kuhn–Tucker vector $\lambda = (\lambda_k) \in R^m$ which satisfies both (6.1) and the complementary slackness condition $\lambda_k f^k(\bar{x}) = 0$ for all $k \in \mathcal{P}$, then

$$\mu(\alpha\epsilon) - \mu(0) \geq -\alpha(\lambda \cdot \epsilon) \quad \text{for all } \alpha \geq 0, \tag{6.11}$$

i.e. the marginal improvement of the optimal value with respect to perturbations in the direction ϵ is bounded below by $-\lambda \cdot \epsilon$. Conversely, if the marginal improvement is bounded below in all directions ϵ, then $\bar{x}$ is a Kuhn–Tucker point. It is interesting to note that in order to verify stability one need only check the perturbation direction $\epsilon = (\epsilon_k)$ with $\epsilon_k = 1$ for all $k \in \mathcal{P}$. Moreover, if Slater's condition is not satisfied and $\epsilon_k < 0$ for all $k \in \mathcal{P}$, then $\mu(\epsilon) = +\infty$ and (6.11) still holds. Gauvin [17] has shown that Slater's condition is equivalent to having a bounded set of Kuhn–Tucker vectors. This is often taken as the definition of stability since we can then allow an arbitrary perturbation vector ϵ and still maintain feasibility. This type of stability is related to stability of perturbations of the feasible set as studied by Robinson [25, 26] and Tuy [29]. In particular, Tuy's notion of stability, given for the abstract program with cone constraints (see also [13, 14]), guarantees the existence of a Kuhn–Tucker vector by requiring that all perturbations in a neighbourhood of the origin be stable. By restricting the perturbations to a subspace containing the range of the feasible set, he is able to weaken the stability (regularity) condition. For example (see [9]) perturbations in the subspace

$$Y = \{\epsilon = (\epsilon_k) : \epsilon_k = 0 \quad \text{for all } k \in \mathcal{P}^=\} \tag{6.12}$$

maintain feasibility and stability. However, $\mathcal{P}^=$ non-empty does not guarantee instability in the sense of (6.11) since a Kuhn–Tucker vector may still exist. In fact, when $\mathcal{P}^= \neq \emptyset$, $\mathcal{P}^b(\bar{x}) = \emptyset$ and $B_{\mathcal{P}(\bar{x})}(\bar{x})$ is closed, then $\bar{x}$ is a regular point in our sense though not in the sense of Tuy. Moreover, under the closure conditions (5.3), we can replace $\mathcal{P}^=$ by $\mathcal{P}^b(\bar{x})$ in (6.12) and still have stable perturbations (see [32] for details). Note that it makes sense to speak of these perturbations as being stable though only the positive ones may maintain feasibility. For example, Zoutendijk [35] suggests that: if Slater's condition fails one should find $\mu(\epsilon)$ with $\epsilon_k > 0$ for all $k \in \mathcal{P}$. This makes sense if $\bar{x}$ is a Kuhn–Tucker point. Otherwise, the marginal improvement of $\mu(\epsilon)$ will be $-\infty$.

7. Regularization

Gould and Tolle have posed the question: "Can the program (P) be regularized by the addition of a finite number of constraints?" Augunwamba [4] has considered the nonconvex, differentiable case and has shown that one can always regularize with the addition of an infinite number of constraints. He has also given necessary and sufficient conditions to insure the number of constraints added may be finite. In this section, we show that one can always regularize (P) at x, by the addition of one (possibly nondifferentiable) constraint. Furthermore, in the case of faithfully convex constraints, we can regularize (P) by the addition (or substitution) of a finite number of linear constraints.

Theorem 7.1. *Suppose that $\bar{x} \in S$, X is a Hilbert space, $\mathcal{P}^b(\bar{x}) \subset \Omega \subset \mathcal{P}^=$, $B_{\mathcal{P}(\bar{x})}(\bar{x})$ is closed and either* conv $D_\Omega^=(\bar{x})$ *is closed or $\Omega = \mathcal{P}^=$. Consider program* (P) *with the additional constraint*

$$f^{m+1}(x) \triangleq \operatorname{dist}((x - \bar{x}), \overline{\operatorname{conv}}\, D_\Omega^=(\bar{x})).$$

Then $\bar{x}$ is a regular point.

Proof. By Lemma 3.1, f^{m+1} is not "badly behaved" at $\bar{x}$ and therefore, $\mathcal{P}^b(\bar{x})$ is not increased by the addition of f^{m+1}. Now, by Theorem 6.1, we need only show that

$$C_{\mathcal{P}(\bar{x}) \cup \{m+1\}}(\bar{x}) \subset \overline{D_{\mathcal{P}^=}^=(\bar{x})}. \tag{7.1}$$

But

$$C_{m+1}(\bar{x}) = \{d \in X : \nabla f^{m+1}(\bar{x}; d) \le 0\} = \overline{\operatorname{conv}}\, D_\Omega^=(\bar{x}) \quad \text{by (3.3).}$$

The inclusion (7.1) now follows from Lemma 4.1(b).

Note that the feasible set remains unchanged after the addition of f^{m+1}. For, let $\bar{S}$ denote the feasible set after the addition. Then

$$\begin{aligned} x \in \bar{S} &\Leftrightarrow x \in S \quad \text{and} \quad x - \bar{x} \in \overline{\operatorname{conv}}\, D_{\Omega}^{\bar{=}}(\bar{x}) \\ &\Leftrightarrow x \in S \quad \text{since } \Omega \subset \mathcal{P}^{=} \quad \text{and} \quad D_{\Omega}^{\bar{=}}(\bar{x}) \subset \overline{\operatorname{conv}}\, D_{\Omega}^{\bar{=}}(\bar{x}). \end{aligned}$$

We have, therefore, regularized the point $\bar{x}$, by the addition of a "redundant" constraint.

Example 7.1. Consider program (P) with the constraints

$$f^k(x) \le 0, \quad k \in \mathcal{P} = \{1, 2\},$$

where f^1 and f^2 are given in Example 5.2. Let $\bar{x} = 0$. Then $\bar{x}$ is not a regular point, since $\mathcal{P}^b(\bar{x}) = \{2\} \ne \emptyset$. Now

$$D_{\mathcal{P}^b(\bar{x})}^{\bar{=}}(\bar{x}) = \{x \in R : x \le 0\}.$$

Therefore adding the redundant constraint

$$f^3(x) = \operatorname{dist}((x - \bar{x}), \overline{\operatorname{conv}}\, D_{\mathcal{P}^b(\bar{x})}^{\bar{=}}(\bar{x})) \le 0$$

or equivalently adding

$$f^3(x) = x \le 0$$

regularizes the point $\bar{x}$.

When the constraints f^k, $k \in \mathcal{P}^{=}$, are faithfully convex, then we know that $D_{\mathcal{P}^=}^{\bar{=}}$ is a closed subspace independent of x. Suppose that $B : Y \to X$ is a linear operator such that $D_{\mathcal{P}^=}^{\bar{=}} = \mathcal{R}(B)$, where Y is a *lcs*. In this case, adding the redundant constraint

$$f^{m+1}(x) = \operatorname{dist}((x - \bar{x}), D_{\mathcal{P}^=}^{\bar{=}})$$

is equivalent to adding the linear constraint

$$x = \bar{x} + By \quad \text{for some } y \in Y.$$

Moreover, we get the following result.

Theorem 7.2. *Let $\bar{x} \in S$ and f^k, $k \in \mathcal{P}^{=}$, be faithfully convex. Suppose that $B : Y \to X$ is a linear operator satisfying*

$$D_{\mathcal{P}^=}^{\bar{=}} = \mathcal{R}(B),$$

where Y is a lcs. Consider the program, in the variable $y \in Y$,

$$(\mathrm{P}_r) \qquad \begin{aligned} &f^0(\bar{x} + By) \to \min \\ &\text{s.t.} \quad f^k(\bar{x} + By) \le 0, \quad k \in \mathcal{P} \setminus \mathcal{P}^{=}. \end{aligned}$$

Then Slater's condition is satisfied for (P_r), *and $y = 0$ is a feasible point of* (P_r). *Moreover, if y^* solves* (P_r), *then $\bar{x} + By^*$ solves* (P).

Proof. The result follows from Theorem 6.1 and the fact that $\mathcal{P}^= = \emptyset$ if and only if Slater's condition holds.

The above result also follows from the characterization of optimality given in [2, 8]. Note that after the substitution, (P_r) has fewer constraints and, as shown in the following example, fewer variables.

Example 7.2. Consider the feasible set S defined by the constraints in Example 6.1. We found that $\mathcal{P}^= = \{1, 4, 5\}$,

$$D^=_{\mathcal{P}^=} = \mathcal{R}\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)$$

and that every $x \in S$ is not regular. By the above theorem, S_r, the feasible set of the regularized program (P_r), is defined by the four constraints in three variables:

$$\begin{aligned} f^2(x) &= e^{-y_1} - 1 \le 0, \\ f^3(x) &= y_2^2 + y_3^2 - 1 \le 0, \\ f^6(x) &= e^{-y_2} - 1 \le 0, \\ f^7(x) &= e^{-y_3} - 1 \le 0. \end{aligned}$$

Note that every $x \in S_r$ is now a regular point and Slater's condition is satisfied.

The above theorem was used in [31] to formulate the Method of Reduction, which first finds a feasible point and then solves program (P) in the case of faithfully convex constraints. Stability of the algorithm can be checked using the WCQ's of the previous section. These and other related results will be presented in a further study.

Acknowledgment

The author would like to thank both Professor Jonathan Borwein, for correcting an error in the original manuscript as well as providing other helpful comments, and the referees, for their very detailed suggestions and aid in shortening the proofs.

References

[1] J. Abadie, "On the Kuhn–Tucker theorem", in: J. Abadie, ed., *Nonlinear Programming* (North-Holland Publishing Co., Amsterdam, 1967).

[2] R.A. Abrams and L. Kerzner, "A simplified test for optimality", *Journal of Optimization Theory and Applications* 25 (1978) 161–170.

[3] K. Arrow, L. Hurwicz and H. Uzawa, "Constraint qualifications in maximization problems", *Naval Research Logistics Quarterly* 8 (1961) 175–191.
[4] C.C. Augunwamba, "Optimality conditions: constraint regularization", *Mathematical Programming* 13 (1977) 38–48.
[5] M.S. Bazaraa, J.J. Goode and C.M. Shetty, "Constraint qualifications revisited", *Management Science* 18 (1972) 567–573.
[6] M.S. Bazaraa, C.M. Shetty, J.J. Goode and M.Z. Nashed, "Nonlinear programming without differentiability in Banach spaces: necessary and sufficient constraint qualification", *Applicable Analysis* 5 (1976) 165–173.
[7] A. Ben-Israel, A. Ben-Tal and S. Zlobec, "Characterization of optimality in convex programming without a constraint qualification", *Journal of Optimization Theory and Applications* 20 (1976) 417–437.
[8] A. Ben-Israel, A. Ben-Tal and S. Zlobec, "Optimality conditions in convex programming", The IX International Symposium on Mathematical Programming (Budapest, Hungary, August 1976).
[9] A. Ben-Tal and A. Ben-Israel, "Characterizations of optimality in convex programming: the nondifferentiable case", *Applicable Analysis*, to appear.
[10] A. Ben-Tal, L. Kerzner and S. Zlobec, "Optimality conditions for convex semi-infinite programming problems", Report (1976).
[11] A. Ben-Tal and S. Zlobec, "A new class of feasible direction methods", Report (1977).
[12] J. Borwein, "Proper efficient points for maximizations with respect to cones", *SIAM Journal of Control and Optimization* 15 (1977) 57–63.
[13] J. Borwein and H. Wolkowicz, "Characterizations of optimality without constraint qualification for the abstract convex program. Part I", Report No. 14, Dalhousie University (Halifax, June 1979).
[14] J. Borwein and H. Wolkowicz, "Characterizations of optimality without constraint qualification for the abstract convex program. Part II: finite dimensional range space", Report No. 19, Dalhousie University (Halifax, July 1979).
[15] F.H. Clarke, "Nondifferentiable functions in optimization", *Applied Mathematics Notes* 2 (Dec. 1976).
[16] V.F. Dem'yanov and V.N. Malozemov, *Introduction to minimax*, Translated from Russian by D. Louvish (John Wiley & Sons, New York, 1974).
[17] J. Gauvin, "A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming", *Mathematical Programming* 12 (1977) 136–138.
[18] A.M. Geoffrion, "Duality in nonlinear programming: a simplified applications-oriented development", *SIAM Review* 13 (1971) 1–37.
[19] I.V. Girsanov, *Lectures on mathematical theory of extremum problems*, Lecture Notes in Economics and Mathematical Systems, No. 67 (Springer-Verlag, Berlin, 1972).
[20] F.J. Gould and J.W. Tolle, "Geometry of optimality conditions and constraint qualifications", *Mathematical Programming* 2 (1972) 1–18.
[21] F.J. Gould and J.W. Tolle, "A necessary and sufficient qualification for constrained optimization", *SIAM Journal of Applied Mathematics* 20 (1971) 164–171.
[22] M. Guignard, "Generalized Kuhn–Tucker conditions for mathematical programming problems in a Banach space", *SIAM Journal of Control* 7 (1969) 232–241.
[23] R.B. Holmes, *A course on optimization and best approximation*, Lecture Notes in Mathematics, No. 257 (Springer-Verlag, Berlin, 1972).
[24] B.N. Pshenichnyi, *Necessary conditions for an extremum* (Marcel Dekker, New York, 1971).
[25] S.M. Robinson, "Stability theory for systems of inequalities. Part I: linear systems", *SIAM Journal of Numerical Analysis* 12 (1975) 754–769.
[26] S.M. Robinson, "Stability theory for systems of inequalities. Part II: differentiable nonlinear systems", *SIAM Journal of Numerical Analysis* 13 (1976) 497–513.
[27] R.T. Rockafellar, *Convex analysis* (Princeton University Press, Princeton, NJ, 1970).
[28] R.T. Rockafellar, "Some convex programs whose duals are linearly constrained", in: J.B. Rosen, O.L. Mangasarian and K. Ritter, eds., *Nonlinear programming* (Academic Press, New York, 1970) 293–322.
[29] H. Tuy, "Stability property of a system of inequalities", *Mathematische Operationsforschung und Statistik, Ser. Optimization* 8 (1977) 27–39.

[30] H. Wolkowicz, "Calculating the cone of directions of constancy", *Journal of Optimization Theory and Applications* 25 (1978) 451–457.
[31] H. Wolkowicz, "Constructive approaches to approximate solutions of operator equations and convex programming", Ph.D. Thesis, McGill University (Montreal, Que., 1978).
[32] H. Wolkowicz, "Shadow prices for an unstable convex program", Report No. 15, Dalhousie University (Halifax, May 1979).
[33] S. Zlobec, *Mathematical Programming,* Lecture notes (McGill University, Montreal, Que., 1977).
[34] S. Zlobec, "Feasible direction methods in the absence of Slater's condition", *Mathematische Operationsforschung und Statistik, Ser. Optimization* 9 (1978) 185–196.
[35] G. Zoutendijk, *Methods of feasible directions: a study in linear and non-linear programming* (Elsevier, New York, 1960).