Mathematical Programming with Data Perturbations II

edited by

Anthony V. Fiacco

Department of Operations Research
The George Washington University
School of Engineering and
Applied Science
Washington, D. C.

MARCEL DEKKER, INC. New York and Basel

CHAPTER 4 OPTIMALITY CONDITIONS AND SHADOW PRICES*

HENRY WOLKOWICZ / The University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

We consider an ordinary convex program that does not necessarily attain its infimum. It is well known that the Kuhn-Tucker multipliers, if they exist, provide shadow prices, or sensitivity coefficients, for the marginal improvement of the optimum with respect to perturbations in the right-hand sides of the constraints. Moreover, the Kuhn-Tucker multipliers exist if and only if this marginal improvement is bounded below for all perturbation directions. This is equivalent to "stability" of the program with respect to all nonnegative right-hand side perturbations. We present a group of optimality conditions dependent on a regularizing set G_1. The "restricted" Kuhn-Tucker multipliers, thus obtained, provide shadow prices for unstable programs. Moreover, a class of perturbations is presented which, at the unit price given by the "restricted" Kuhn-Tucker multipliers, are never economical to buy, i.e., we find a stable subset of perturbations.

*Research partially supported by NSERC A3388.

I. INTRODUCTION

In this paper we consider the ordinary convex program

$$\begin{aligned} &\mu = \mathit{inf}\ p(x) \\ &\mathit{subject\ to}\ g^k(x) \leq 0 \qquad k \in P = \{1,\dots,m\} \\ &x \in \Omega \end{aligned} \tag{P}$$

where $p, g^k : U \to R$ are convex functions on U, $U \subset R^n$ is some open set containing Ω, and Ω is a convex set.

If some constraint qualification, e.g., Slater's condition, holds for (P), then

$$\mu = \inf\{p(x) + \sum \lambda_k g^k(x) : x \in \Omega\} \tag{1}$$

for some Kuhn-Tucker multipliers $\lambda_k \geq 0$. The λ_k provide shadow prices for (P), i.e., if

$$\mu(\varepsilon) = \inf\{p(x) : g^k(x) \leq \varepsilon_k,\ k \in P,\ x \in \Omega\} \tag{2}$$

then, at the unit prices λ_k, no perturbation $\varepsilon = (\varepsilon_k)$ whatsoever is economical to buy. Moreover, this equilibrium situation characterizes the Kuhn-Tucker vector $\lambda = (\lambda_k)$, see, e.g., [11].

In the absence of any constraint qualification, no Kuhn-Tucker vector may exist. Characterizations of optimality without any constraint qualification, using feasible directions, have been given in, e.g., [1,2,3,4]. Stronger optimality conditions were given in [13,14,15,16]. These conditions all assumed that the infimum is attained and that $\Omega = R^n$.

In this paper we present several optimality conditions without constraint qualification of the type (1). In these conditions, Ω is replaced by $\Omega \cap G$, where G is some appropriate convex set. These results generalize the above mentioned characterizations, in the sense that the infimum need not be attained and Ω may be a proper, not necessarily polyhedral, subset of R^n. The proofs for these conditions are unified and depend on the characterization of optimality (with constraint qualification) given in Lemma 2 below. The unifying key to the proofs and method of choosing the sets G is that the equality set of constraints, denoted $P^=$, are affine on $\Omega \cap G$. (A generalization of the characterization of optimality in [1,2,4], to the case when the infimum may not be attained and $\Omega \neq R^n$, was first given in [5]. This generalization followed from a characterization of optimality for the abstract convex program.)

For these optimality conditions without constraint qualification, we find a set of perturbations $\varepsilon(G)$ which, at the unit prices λ_k, are never economical to buy, i.e., a stable set of perturbations. (Note that, in the case that the infimum is attained, the indices k such that the perturbation $t\varepsilon_k$ is always economical to buy for some $t > 0$, no matter what the cost, were found in [14]; see Section IV below.) We conclude with two examples.

II. PRELIMINARIES

Consider the ordinary convex program (P). Let $g(x) = (g^k(x))$ denote the vector of constraints in R^m. We assume that the *feasible set*

$$F = \{x \in \Omega : g(x) \leq 0\} \neq \emptyset \tag{3}$$

The inequalities $g(x) \leq 0$ and $g(x) < 0$ are taken component-wise.

We can now state the "standard" Karush-Kuhn-Tucker theorem, see, e.g., [11]. Note that a constraint qualification is a condition that guarantees the existence of a Kuhn-Tucker vector.

THEOREM 1 Suppose that some constraint qualification holds for (P), e.g.,

there exists $\hat{x} \in \Omega$ such that $g(\hat{x}) < 0$ (Slater's condition)

Then

$$\mu = \inf\{p(x) + \lambda g(x) : x \in \Omega\} \tag{4}$$

for some Kuhn-Tucker vector $\lambda = (\lambda_k) \geq 0$. Moreover, if $\mu = p(x^*)$ for some x^* in F, then

$$\lambda g(x^*) = 0 \qquad \text{(complementary slackness)} \tag{5}$$

and (4) and (5) characterize optimality of x^* in F.

The inner product $\lambda g(x)$ of the two vectors λ and $g(x)$ is stated by juxtaposition.

Now if $f : U \to R$ is a convex function on U and $x \in U$, let

$$\partial f(x) = \{\phi \in R^n : \phi(y - x) \leq f(y) - f(x), \text{ for all } y \text{ in } U\} \tag{6}$$

denote the *subdifferential of* f *at* x. For $K \subset R^n$, let

$$K^+ = \{\phi \in R^n : \phi y \geq 0, \text{ for all } y \text{ in } K\} \tag{7}$$

be the *nonnegative polar cone* of K. Then we have the following optimality characterization for a convex function f on U.

LEMMA 1 (e.g., [10, p. 87]). Suppose that $x^* \in \Omega$. Then

$$f(x^*) = \inf\{f(x) : x \in \Omega\}$$

if and only if

$$\partial f(x^*) \cap (\Omega - x^*)^+ \neq \phi \tag{8}$$

If some constraint qualification holds for (P), this result and Theorem 1 yield:

$$\begin{aligned}&x^* \text{ in } F \text{ is optimal for (P) if and only if}\\ &0 \in \partial p(x^*) + \partial\lambda g(x^*) - (\Omega - x^*)^+\\ &\text{for some } \lambda = (\lambda_k) \geq 0 \text{ with } \lambda g(x^*) = 0\end{aligned} \tag{9}$$

Note that p and λg are continuous on U, since they are convex and finite on the open set U. Thus

$$\partial(p(x^*) + \lambda g(x^*)) = \partial p(x^*) + \partial\lambda g(x^*)$$

The *directional derivative* of the convex function f at x in U in the direction d is

$$\nabla f(x;d) = \lim_{t \downarrow 0} \frac{f(x+td) - f(x)}{t} \tag{10}$$

Let

$$\mathcal{P}(x) = \{k \in \mathcal{P} : g^k(x) = 0\} \tag{11}$$

be the set of binding constraints at x in F. The set of *implicit equality constraints*, e.g., [1], is

$$\mathcal{P}^= = \{k \in \mathcal{P} : g^k(x) = 0, \text{ for all } x \text{ in } F\} \tag{12}$$

We set

$$\mathcal{P}^<(x) = \mathcal{P}(x) \setminus \mathcal{P}^= \tag{13}$$

$$C(x) = \{d \in R^n : \nabla g^k(x;d) \leq 0, \text{ for all } k \in \mathcal{P}(x)\} \tag{14}$$

is the *linearizing cone* at x, while

$$B(x) = \text{cone} \bigcup_{k \in \mathcal{P}(x)} \partial g^k(x)$$

is the *cone of subgradients* at x, where cone denotes the generated convex cone.

$$T(F,x) = \overline{\text{cone}}(F - x) \tag{15}$$

is the tangent cone of the convex feasible set F at x, where $\overline{\cdot}$ denotes closure.

For the *perturbed program*

$$\mu(\varepsilon) = \inf\{p(x) : g(x) \leq \varepsilon,\ x \in \Omega\} \qquad (P_\varepsilon)$$

$\mu(\varepsilon)$ is called the *perturbation function*. We let $F_\varepsilon = \{x : g(x) \leq \varepsilon,\ x \in \Omega\}$ denote the *perturbed feasible set*.

The *cone of directions of constancy* of a convex function f at x is (see, e.g., [2])

$$D_f^{=}(x) = \{d \in R^n : \text{there exists } \bar{\alpha} > 0 \text{ with } f(x + \alpha d) = f(x), \text{ for all } 0 < \alpha \leq \bar{\alpha}\} \qquad (16)$$

We let $D_k^{=}(x) = D_{g_k}^{=}(x)$ and $D_K^{=}(x) = \bigcap_{k \in K} D_k^{=}(x)$, where $K \subset P$. The *cones of directions of increase, decrease, and nonincrease*, denoted $D_f^{>}(x)$, $D_f^{<}(x)$, and $D_f^{\leq}(x)$, are similarly defined. We also let

$$D_f^{a}(x) = \{d \in R^n : \text{there exists } \bar{\alpha} > 0, \text{ with } f(x + \alpha d) \text{ affine}, \text{ for all } 0 < \alpha \leq \bar{\alpha}\}$$

denote the *cone of affine directions*; see [16].

III. RESTRICTED KUHN-TUCKER VECTORS

In this section we present several optimality conditions for (P) of the type (1), with Ω replaced by $\Omega \cap G$. These results extend the conditions given in [1,2,3,4] and [13,14,15,16] in the sense that the infimum need not be attained and Ω may be a proper subset of R^n. The proofs given here are unified and are based on Lemma 2 below. The idea is to find a regularizing set G, in order to satisfy the generalized Slater condition given in the lemma.

For each choice of the set G and corresponding vector λ which satisfies (1), with Ω replaced by $\Omega \cap G$, we find a set of stable perturbations; see Theorem 3 below.

We call $\lambda = (\lambda_k) \geq 0$ a *restricted Kuhn-Tucker vector* with respect to the convex set G if

$$\mu \leq \inf\{p(x) + \lambda g(x) : x \in \Omega_G = \Omega \cap G\} \qquad (17)$$

If the infimum for (P) is attained at x* in F, i.e., $\mu = p(x^*)$, then we can trivially choose $\lambda = 0$ and $G = \{x^*\}$. Or, in the case that the infimum is not attained, we can choose $\lambda = 0$ and $G = F$, the feasible set. Now consider the perturbation $\varepsilon = (\varepsilon_k)$ and suppose that (17) holds and

$$F_\varepsilon \subset \Omega_G \tag{18}$$

Then by (17),

$$\begin{aligned} p(x) - \mu &\geq -\lambda g(x) && \text{for all } x \text{ in } \Omega_G \\ &\geq -\lambda\varepsilon && \text{for all } x \text{ in } F_\varepsilon \text{ by (18)} \end{aligned}$$

Taking the infimum over x yields

$$\mu(\varepsilon) - \mu \geq -\lambda\varepsilon \qquad \text{if (18) holds} \tag{19}$$

Thus the restricted Kuhn-Tucker vector λ provides a lower bound for the marginal improvement of the optimal value with respect to the perturbation ε if $F_\varepsilon \subset \Omega_G$. Thus the program is stable, or equivalently λ is a Kuhn-Tucker vector, if we can choose $G \supset \Omega$. Moreover, the program will be stable with respect to a larger set of perturbations ε if we can choose the set G larger. This motivates the search for larger sets G in (17) (or stronger optimality conditions, e.g., Gould and Tolle [8] and Guignard [9]). We will now present several choices for the set G in (17). First we need the following lemma. Note that a *polyhedral* function is the maximum of a finite number of affine functions.

LEMMA 2 [6] Consider the ordinary convex program (P). Suppose that Ω is a polyhedral set and that there exists an $x_1 \in \Omega$ such that $g^k(x_1) \leq 0$, $k = 1,\ldots,m$, with strict inequality if g^k is not polyhedral (on Ω). Then

$$\mu = \inf\{p(x) + \lambda g(x) : x \in \Omega\} \tag{20}$$

for some $\lambda = (\lambda_k) \geq 0$. Moreover, if $\mu = p(a)$ for some $a \in F$, then

$$\lambda g(a) = 0 \tag{21}$$

and (20) and (21) characterize optimality of a.

REMARK 1. If Ω is not polyhedral, then the above result holds if either $x_1 \in \operatorname{ri}\Omega$, the relative interior of Ω, or if the strict inequality, $g^k(x_1) < 0$, holds for all k in P; see, e.g., [11].

Following Rockafellar [12], we say that a function $f : U \to R$ is *faithfully convex* (on U) if f is convex (on U) and f is affine on a line segment (in U) only if it is affine on the whole line (in U) containing that segment. Analytic convex and strictly convex functions are two examples of faithfully convex functions. If f is faithfully convex on U, then it can be shown that the cone of directions of constancy $D_f^=(x)$ is a subspace independent of the point x in U (see, e.g., [2]). We can now derive various choices for the set G in (17).

THEOREM 2 Suppose that Ω is polyhedral, g^k is faithfully convex (on U) for all $k \in P^=$, and $h = \sum_{k \in P^=} \alpha_k g^k$, where $\alpha_k \geq 0$ with $\alpha_k > 0$ if g^k is not affine (on U). Let $\hat{x} \in F$. Then there exist convex sets G such that

$$\mu = \inf\{p(x) + \lambda g(x) : x \in \Omega_G = \Omega \cap G\} \tag{22}$$

for some $\lambda \geq 0$ and, if $\mu = p(x^*)$ for some x^* in F, then

$$\lambda g(x^*) = 0 \tag{23}$$

and (22) and (23) characterize optimality of x^*. Possible choices of the convex set G in (22) are:

(a) $G_1 = \{x : g^k(x) = 0 \text{ for all } k \in P^=\}$

(b) $G_2 = \hat{x} + D_f^=(\hat{x})$, where $f \triangleq \max_{k \in P^=} g^k$

(c) $G_3 = \hat{x} + D_{P^=}^=(\hat{x})$

(d) $G_4 = \hat{x} + D_h^=(\hat{x})$

(e) $G_5 = \hat{x} + D_{P^=}^a(\hat{x})$

(f) $G_6 = \hat{x} + D_h^a(\hat{x})$

(Note that we can allow λ_k to be arbitrary for all k such that $g^k(x) = 0$ on $\Omega \cap G_i$.)

PROOF. (a): It can be shown that $G_1 = G_3$; see, e.g., [2]. Now see the proof of (c) below. (b) and (c): First let us show that $G_2 = G_3$. That $D_{P^=}^=(x) \subset D_f^=(x)$ is clear. Now suppose that $d \in D_f^=(x) \setminus D_{P^=}^=(x)$. Then, by definition of f, $d \in D_{P^=}^{\leq}(x) \cap D_k^<(x)$ for some $k \in P^=$. It can be shown (see [2]) that there exists

$$\hat{d} \in D_{P^=}^=(\hat{x}) \cap D_{P^<(\hat{x})}^<(\hat{x}) \tag{24}$$

(The proof of (24) remains the same even though $\Omega \neq R^n$.) Let

$$d_\lambda = \lambda d + (1 - \lambda)\hat{d} \tag{25}$$

Then for sufficiently small $\lambda > 0$, we see that

$$d_\lambda \in D_{P(\hat{x})}^{\leq}(\hat{x}) \cap D_k^<(\hat{x}) \tag{26}$$

which contradicts the definition of $P^=$. Thus

$$D_f^=(\hat{x}) = D_{P^=}^=(\hat{x})$$

Now it is clear that $F \subset \hat{x} + D_{P^=}^=(\hat{x})$. Thus an equivalent program to (P) is

$$\mu = \inf\{p(x) : g(x) \leq 0, x \in \Omega \cap G_3\} \tag{27}$$

It is now sufficient to show that we can satisfy the hypothesis of Lemma 2 with Ω replaced by $\Omega \cap G_3$. Now be definition of $P^=$, for each $k \in P \backslash P^=$ there exists $x^k \in F$ such that $g^k(x^k) < 0$. Let t equal the cardinality of $P \backslash P^=$. Then, by the convexity of F, we see that

$$x_1 = \sum_{k \in P \backslash P^=} \frac{1}{t} x^k \in F$$

and

$$g^k(x_1) < 0, \ k \in P \backslash P^=; \quad g^k(x_1) = 0, \ k \in P^=; \quad x_1 \in \Omega \tag{28}$$

Moreover, this implies that

$$x_1 - \hat{x} \in D_{P^=}^=(\hat{x}) \tag{29}$$

Thus x_1 satisfies the constraint qualification in Lemma 2 if we show that

$$g^k \text{ is affine on } G_3 \cap \Omega \qquad \text{for all } k \in P^= \tag{30}$$

In fact, by the faithfully convex assumption, it is clear that g^k is 0 on $G_3 \cap \Omega$, for all $k \in P^=$. This proves that G_3, and so also G_2, is a possible choice for G in (22). Note that $\Omega \cap G_2$ is polyhedral since Ω is, by assumption, and G_2 is an affine subspace, by faithful convexity.

Following the details of the proof above, we see that to show that G_i, $i = 4,5,6$, are possible choices for G in (22) we need only show that

$$g^k \text{ is affine on } \Omega \cap G_i \qquad \text{for all } k \in P^=, \ i = 4,5,6 \tag{31}$$

(d): Let $d \in D_h^=$, $k_0 \in P^=$, and suppose g^{k_0} is not affine (on Ω). Then $\alpha_{k_0} > 0$ and

$$-\alpha_{k_0} g^{k_0}(\hat{x} + \alpha d) = \sum_{k \in P^= \backslash \{k_0\}} \alpha_k g^k(\hat{x} + \alpha d) \tag{32}$$

Since a nonnegative linear combination of convex functions is convex, this shows that both $\pm g^{k_0}(\hat{x} + \alpha d)$ is a convex function of α. Thus g^k is in fact an affine function of α. This shows that g^k is affine on $\Omega \cap G_4$ as desired.

The result with G_5 and G_6 follows similarly by allowing an affine term on the right-hand side of (32).

REMARK 2. The characterization of optimality using G_1 holds without the polydedrality and faithful convexity assumptions; see [5]. This extends the characterization of optimality given in [1,2,4,17], which requires that the infimum for (P) be attained and that $\Omega = \mathbb{R}^n$. Using Lemma 1, we get

$$\begin{aligned}&x^* \text{ (in } F) \text{ is optimal for (P) if and only if}\\ &0 \in \partial p(x^*) + \partial \lambda g(x^*) - (\Omega_{1_{G_1}} - x^*)^+\\ &\text{for some } \lambda = (\lambda_k) \geq 0 \text{ with } \lambda g(x^*) = 0\end{aligned}$$

It can be shown (see [5]) that

$$(\Omega_{1_{G_1}} - x^*)^+ = \left(D^{=}_{P^{=}}(x^*) \cap \text{cone}(\Omega - x^*)\right)^+$$

and that λ can be chosen with $\lambda_k = 0$ if $k \in P^{=}$. If we in fact fix $\Omega = \mathbb{R}^n$ and choose λ with $\lambda_k = 0$ for $k \in P^{=}$, then (32) becomes the result in [4]. Note that this result is a special case of the generalized Kuhn-Tucker conditions given by Guignard [9].

We can similarly recover the results obtained in [13,14,15,16]. Moreover, one can replace the assumption of faithful convexity (and affinity) by piecewise faithful convexity (resp. polyhedrality), i.e., that g^k, $k \in P^{=}$, is the maximum of a finite number of faithfully convex functions. For, working with the equivalent program to (P), obtained by replacing the piecewise faithfully convex functions by the faithfully convex functions from which the maximum is taken, gives rise to a minorant for the Lagrangian, i.e.,

$$\begin{aligned}\mu = \inf&\left\{p(x) + \sum_{k,t} \lambda_{k,t} g^{k,t}(x) : x \in \Omega \cap G\right\}\\ &\leq \inf\{p(x) + \lambda g(x) : x \in \Omega \cap G\}\end{aligned}$$

if $\lambda = (\lambda_k)$, $\lambda_k = \sum_t \lambda_{k,t}$, and $g^k = \max_t g^{k,t}$.†

We now define the following class of "stable" perturbations in $\mathbb{R}^m$, with respect to a convex set G in $\mathbb{R}^n$,

†This argument is due to Professor Jon Borwein.

$$\varepsilon(G) = \left\{\varepsilon \in R^m : \mu(\varepsilon) = \inf\{p(x) : g(x) \leq \varepsilon,\ x \in \Omega \cap G\}\right\} \tag{33}$$

i.e., this is the set of perturbations for which the perturbation function does not change when the (regularizing) set G is added. We now have the following property of a restricted Kuhn-Tucker vector in terms of equilibrium prices for the perturbation.

THEOREM 3 When μ is finite, then $\lambda^* = (\lambda_k^*) \geq 0$ is a restricted Kuhn-Tucker vector, with respect to the convex set G, implies that at the unit price λ_k^* for a perturbation ε_k, no perturbation in $\varepsilon(G)$ is worth buying.

PROOF. That no perturbation in $\varepsilon(G)$ is worth buying is equivalent to the inequality

$$\mu(\varepsilon) + \lambda^*\varepsilon \geq \mu \qquad \text{for all } \varepsilon \in \varepsilon(G) \tag{34}$$

For any perturbation ε, the minimum value in the perturbed program (P_ε) plus the cost of buying the perturbation ε, at the unit prices given by the vector $v = (v_k)$, is

$$\mu(\varepsilon) + v\varepsilon$$

Now, for any m-vector $\lambda \geq 0$,

$$\inf\{\mu(\varepsilon) + \lambda\varepsilon : \varepsilon \in \varepsilon(G)\} = \inf\{p(x) + \lambda\varepsilon : \varepsilon \in \varepsilon(G),\ x \in F_\varepsilon\}$$

$$= \inf\{p(x) + \lambda\varepsilon : \varepsilon \in \varepsilon(G),\ x \in \Omega,\ g(x) \leq \varepsilon\}$$

$$= \inf\{p(x) + \lambda\varepsilon : \varepsilon \in \varepsilon(G),\ x \in \Omega_G,\ g(x) \leq \varepsilon\} \quad \text{by definition of } \varepsilon(G)$$

$$= \inf\{p(x) + \lambda g(x) : \varepsilon \in \varepsilon(G),\ x \in \Omega_G,\ g(x) \leq \varepsilon\} \tag{35}$$

[since $\lambda \geq 0$ and $\varepsilon_1 \in \varepsilon(G)$, $\varepsilon_2 \leq \varepsilon_1$ implies $\varepsilon_2 \in \varepsilon(G)$]

$$\geq \inf\{p(x) + \lambda g(x) : x \in \Omega_G\}$$

$$\geq \mu \qquad \text{if and only if (17) holds, i.e., if } \lambda = \lambda^*$$

An alternative proof of the above is

$$\begin{aligned} p(x) - \mu &\geq -\lambda^* g(x) && \text{for all } x \text{ in } \Omega \cap G, \text{ by assumption} \\ &\geq -\lambda^*\varepsilon && \text{if } x \in \Omega \cap G \text{ and } g(x) \leq \varepsilon \end{aligned}$$

Thus

$$p(x) - \mu \geq -\lambda^*\varepsilon \qquad \text{for all } x \text{ in } F_\varepsilon \text{ if } \varepsilon \in \varepsilon(G)$$

Taking the infimum over x on both sides yields

$$\mu(\varepsilon) + \lambda^*\varepsilon \geq \mu$$

Note that if some constraint qualification holds then λ is a Kuhn-Tucker vector if and only if, at the unit prices λ_k, no perturbation whatsoever is economical to buy; see [11, p. 277]. It is conjectured that "if and only if" holds in the above theorem as well. For this to hold, it remains to show that equality holds in (35). Otherwise, we can find a larger set of perturbations than $\varepsilon(G)$.

COROLLARY 1 Suppose that the perturbation ε satisfies

$$F_\varepsilon \subset \Omega \cap G \tag{36}$$

and λ is a restricted Kuhn-Tucker vector satisfying (17). Then ε is not worth buying at the unit prices given by $\lambda = (\lambda_k)$.

PROOF. It is clear that (36) implies $\varepsilon \in \varepsilon(G)$.

The corresponding *restricted Lagrangian dual program* with respect to the set G is

$$\nu = \sup_{\lambda \geq 0} \inf_{x \in \Omega \cap G} \{p(x) + \lambda g(x)\} \tag{D_G}$$

Note that if λ is a solution of (17), then $\nu = \mu$ (no duality gap) and λ is also a solution of (D_G).

IV. PROGRAMS WITH ATTAINED INFIMUM

Let us now assume that x^* is in F, $\mu = p(x^*)$, and both p and g are continuous at x^*. We now look for optimality conditions of the type

$$\begin{aligned} &x^* \text{ in } F \text{ is optimal if and only if} \\ &0 \in \partial p(x^*) + \partial \lambda g(x^*) - (\Omega_G - x^*)^+ \qquad \text{for some } \lambda \geq 0 \text{ with } \lambda g(x^*) = 0 \end{aligned} \tag{37}$$

where $\Omega_G = \Omega \cap G$. By applying Lemma 1, these conditions can be reformulated in the same form as the conditions given in Theorem 2.

We now present some choices of G in (37) different and possibly larger than those which can be obtained from Theorem 2. These sets G can be found by satisfying the equation (x^* is an optimal solution)

$$T^+(F,x^*) = -B(x^*) + (\Omega_G - x^*)^+$$

In the differentiable case, $B(x^*)$ can be replaced by $C(x^*)$. (See [8,9,13] for more details.) Note that larger sets G in (17), and so "stronger" optimality conditions, correspond to "smaller" sets $(\Omega_G - x^*)$ in (37).

Now let

$$P^b(x) = \left\{k \in P^= : C(x) \cap D_k^>(x) \setminus \overline{D_{P^=}^=} \neq \emptyset\right\} \tag{38}$$

This is the set of *badly behaved constraints* at x (see [13]), i.e., the set of constraints which create problems in the Kuhn-Tucker theory. It can be shown [13] that

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

is a weakest constraint qualification at x, i.e., is a necessary and sufficient condition for the Kuhn-Tucker theory to hold at x. We can introduce the objective function into this definition by setting

$$P_p^b(x) = \left\{k \in P^= : C(x) \cap D_k^>(x) \cap D_p^<(x) \neq \emptyset\right\} \tag{39}$$

For simplicity we assume that $\Omega = R^n$ and g^k, $k \in P^=$, is faithfully convex and differentiable. Let $x \in F$ and

$$P^b(x) \subset K \subset P^=$$

Then we can choose (see [14])

$$G = \left(D_K^=(x)\right)^+$$

in (37). This is also true with $P^b(x)$ replaced by $P_p^b(x)$; see [14]. It can also be shown that $D_K^=(x)$ can be replaced by $D_K^a(x)$, $D_f^=(x)$, or $D_f^a(x)$, where $f = \sum_{k\in K} \alpha_k g^k$ and α_k are any positive scalars.

Moreover, it was shown in [14] that, discounting redundancies in the constraints, the shadow price λ_k corresponding to $k \in P_f^b(x^*)$ is essentially infinite, i.e., for $k \in P_f^b(x^*)$, small amounts of the perturbation ε_k are economical to buy no matter how high the unit cost.

EXAMPLE 1. Consider the (linear) program

$$\begin{aligned}
\mu = \mathit{inf}\ p(x) &= cx = x_1 + 2x_2 + 3x_3 \\
\text{s.t. } g^1(x) &= a^1x - b_1 = x_1 + x_2 + x_3 - 1 \leq 0 \\
g^2(x) &= a^2x - b_2 = -x_1 - x_2 - x_3 + 1 \leq 0 \\
g^3(x) &= a^3x - b_3 = 2x_1 + 2x_2 - 1 \leq 0 \\
g^4(x) &= a^4x - b_4 = -2x_1 - 2x_2 + 1 \leq 0 \\
g^5(x) &= a^5x - b_5 = -x_2 - x_3 + 1 \leq 0 \qquad x \geq 0
\end{aligned} \tag{40}$$

Then $P = P^{=} = \{1,2,\ldots,5\}$; $F = \{x_1 = 0,\ x_2 = x_3 = 1/2\}$; $\mu = 5/2$; and $G_1 = G_2 = G_3 = F$. Suppose we choose G_1 in Theorem 2. Since the constraints g^i are all identically 0 on G_1, any $\lambda \geq 0$ solves (17) with $G = G_1$. However, not every $\lambda \geq 0$ can be interpreted as a shadow price for the original program (P), e.g., $\lambda = 0$ is clearly not a shadow price for the perturbation $\varepsilon_k = 1$, for all $k \in P$. In addition, the corresponding restricted dual is identical to the primal program (40). Thus we get no duality information when using the regularizing set G_1.

Now let $\alpha_k = 1$ and $h(x) = \sum_{k\in P^{=}} \alpha_k g^k(x) = -x_2 - x_3 + 1$. Then

$$G_4 = \{x : x_2 + x_3 = 1\}$$

The corresponding restricted dual is

$$\mu = \sup_{\lambda\geq 0} \inf_{\substack{x\geq 0\\ x_2+x_3=1}} \left\{cx + \sum_{k=1}^{5} \lambda_k(a^k x - b_k)\right\}$$

$$= \sup_{\lambda\geq 0} \inf_{\substack{x\geq 0\\ x_2+x_3=1}} \left\{cx + \sum_{k=1}^{4} \lambda_k(a^k x - b_k)\right\} \qquad \text{since } g^5 \equiv 0 \text{ on } G_4$$

$$= \sup_{\lambda\geq 0} \inf_{\substack{x\geq 0\\ x_2+x_3=1}} \left\{\left(c + \sum_{k=1}^{4} \lambda_k a^k\right)x - \sum_{k=1}^{4} \lambda_k b_k\right\}$$

$$= \sup_{\substack{\lambda\geq 0\\ 1+\lambda_1-\lambda_2+2\lambda_3-2\lambda_4\geq 0}} \inf_{\substack{x\geq 0\\ x_2+x_3=1}} \left\{(2 + \lambda_1 - \lambda_2 + 2\lambda_3 - 2\lambda_4)x_2 + (3 + \lambda_1 - \lambda_2)x_3 - \sum_{k=1}^{4} \lambda_k b_k\right\}$$

[since the infimum is $-\infty$ if $1 + \lambda_1 - \lambda_2 + 2\lambda_3 - 2\lambda_4 < 0$]

$$= \sup_{1+\alpha+2\beta\geq 0} \inf_{\substack{x\geq 0\\ x_2+x_3=1}} \left\{(2 + \alpha + 2\beta)x_2 + (3 + \alpha)x_3 - (\alpha + \beta)\right\}$$

$$= 2\,\frac{1}{2} \qquad \text{when } \beta = \frac{1}{2}$$

Thus $\alpha \geq -2$ and $\beta = 1/2$, i.e.,

$$\lambda_1 - \lambda_2 \geq -2; \quad \lambda_3 - \lambda_4 = \frac{1}{2}; \quad \lambda_i \geq 0$$

are the solutions of the restricted dual program. The perturbations $\varepsilon = (\varepsilon_k)$ with $\varepsilon_4, \varepsilon_5 \geq 0$ and $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 0$ are in $\varepsilon(G_4)$, since G_4 equals the set of points x which satisfy the constraints g^1, g^2, g^5. In particular, we get that $\nabla\mu(0;e_3) \geq -(1/2)$ and $\nabla\mu(0;e_4) \geq 0$, where e_3 and e_4 are the third and fourth unit vectors, respectively, i.e., we get shadow prices for the third and fourth constraints.

Since the constraints are affine, we have $G_5 = G_6 = R^n$. The corresponding dual reduces to the usual linear programming dual,

$$\begin{cases} \mu = \sup -\lambda b \\ \text{s.t.}\ A^t\lambda - c^t \geq 0 \\ \qquad \lambda \geq 0 \end{cases}$$

where A is the matrix with rows a^i and $\cdot^t$ denotes transpose. Two basic optimal solutions are $(0,2,\frac{1}{2},0,1)$ and $(0,0,\frac{1}{2},0,3)$. This yields $\nabla\mu(0;e_3) \geq -(1/2)$ and $\nabla\mu(0;e_4) \geq 0$, as before.

The example above shows how using larger choices of G in (17) yields more dual information. This is the case for nonlinear problems as well, as the following example illustrates.

EXAMPLE 2. Consider the program

$$\begin{cases} \mu = \inf p(x) = x_2 \\ \text{s.t.}\ g^1(x) = x_1^2 + x_2^2 - 1 \leq 0 \\ \qquad g^2(x) = -x_1 \qquad + 1 \leq 0 \end{cases}$$

Then $P = P^= = \{1,2\}$; $F = \{x_1 = 1, x_2 = 0\} = x^*$; $\mu = 0$; and $G_1 = G_2 = \ldots = G_6 = F$. Choosing G_1 in Theorem 2 yields no information. Thus even though one of the constraints is affine, we cannot obtain any shadow prices. In fact, $\nabla\mu(0;e_i) = -\infty$ for the unit vectors e_1, e_2. This points out the fact that affine constraints cannot always be ignored. Note that $P^b(x^*) = P^b_p(x^*) = \{1\}$ and $D_1^=(x^*) = D_1^a(x^*) = \{0\}$. Thus using only the badly behaved constraints does not help. However, suppose that we change the objective function to be $p(x) = x_1$. Then G_i, $i = 1,\ldots,6$, is unchanged. But now $P^b_p(x^*) = \emptyset$. We can therefore set $G = 0$ and get

$$0 = \begin{pmatrix}1\\0\end{pmatrix} + \lambda_1\begin{pmatrix}2\\0\end{pmatrix} + \lambda_2\begin{pmatrix}-1\\0\end{pmatrix} \qquad \lambda_i \geq 0$$

This yields $\nabla\mu'(0;e_1) \geq 0$; $\nabla\mu'(0;e_2) \geq -1$.

REFERENCES

1. R. A. Abrams and L. Kerzner. A simplified test for optimality, *Journal of Optimization Theory and Applications*, Vol. 25, pp. 161-170 (1978).

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

3. A. Ben-Israel, A. Ben-Tal, and S. Zlobec. Characterizations of optimality in convex programming without a constraint qualification, *Journal of Optimization Theory and Applications*, Vol. 20, pp. 417-437 (1976).

4. A. Ben-Tal and A. Ben-Israel. Characterizations of optimality in convex programming: the nondifferentiable case, *Applicable Analysis*. pp. 137-156 (1979).

5. J. Borwein and H. Wolkowicz. Characterizations of optimality for the abstract convex program with finite dimensional range, *Journal of the Australian Mathematical Society*, Vol. 30, pp. 390-411 (1981).

6. J. Borwein and H. Wolkowicz. Cone-convex programming: stability and affine constraint functions, in *Generalized Concavity in Optimization and Economics* (S. Schaible and W. T. Ziemba, eds.), pp. 379-397, Academic Press, New York, 1981.

7. A. M. Geoffrion. Duality in nonlinear programming: A simplified applications-oriented development, *SIAM Review*, Vol. 13, pp. 1-37 (1971).

8. F. J. Gould and J. W. Tolle. Geometry of optimality conditions and constraint qualification, *Mathematical Programming*, Vol. 2, pp. 1-18 (1972).

9. M. Guignard. Generalized Kuhn-Tucker conditions for mathematical programming problems in Banach space, *SIAM Journal of Control*, Vol. 7, pp. 232-241 (1969).

10. R. B. Holmes. *Geometric Functional Analysis and Its Applications*. Springer-Verlag, New York, 1975.

11. R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

12. R. T. Rockafellar. Some convex programs whose duals are linearly constrained, in *Nonlinear Programming* (J. B. Rosen, O. L. Mangasarian, and K. Ritter, eds.), pp. 293-322, Academic Press, New York, 1970.

13. H. Wolkowicz. Geometry of optimality conditions and constraint qualifications: The convex case, *Mathematical Programming*, Vol. 19, pp. 32-60 (1980).

14. H. Wolkowicz. Shadow prices for an unstable convex program, *Utilitas Mathematica*, Vol. 18, pp. 119-139 (1980).

15. H. Wolkowicz. A strengthened test for optimality, *Journal of Optimization Theory and Applications*, Vol. 35, pp. 497-515 (1981).

16. H. Wolkowicz. Optimality and the cone of affine directions, Research Report, The University of Alberta (1980).

17. S. Zlobec and A. Ben-Israel. Duality in convex programming: A linearization approach, *Mathematische Operationsforsch. Statist. Ser. Opt.*, Vol. 10, pp. 171-178 (1979).