JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 33, No. 4, APRIL 1981

A Limiting Lagrangian for Infinitely Constrained Convex Optimization in R^n[1,2]

R. G. JEROSLOW[3]

Communicated by O. L. Mangasarian

Abstract. For convex optimization in R^n, we show how a minor modification of the usual Lagrangian function (unlike that of the augmented Lagrangians), plus a limiting operation, allows one to close duality gaps even in the absence of a Kuhn–Tucker vector [see the introductory discussion, and see the discussion in Section 4 regarding Eq. (2)]. The cardinality of the convex constraining functions can be arbitrary (finite, countable, or uncountable).

In fact, our main result (Theorem 4.3) reveals much finer detail concerning our limiting Lagrangian. There are affine minorants (for any value $0<\theta\leq 1$ of the limiting parameter θ) of the given convex functions, plus an affine form nonpositive on K, for which a general linear inequality holds on R^n. After substantial weakening, this inequality leads to the conclusions of the previous paragraph.

This work is motivated by, and is a direct outgrowth of, research carried out jointly with R. J. Duffin.

Key Words. Lagrangians, nonlinear programming, Kuhn–Tucker theory, convex optimization.

1. Introduction

We consider convex programs

$$\begin{aligned} &(\mathrm{P}) && \inf f_0(x), && \\ &\text{subject to} && f_h(x)\leq 0, && h\in H, \qquad (1)\\ & && x\in K, && \end{aligned}$$

[1] This research was supported by NSF Grant No. GP-37510X1 and ONR Contract No. N00014-75-C0621, NR-047-048. This paper was presented at "Constructive Approaches to Mathematical Models," a symposium in honor of R. J. Duffin, Pittsburgh, Pennsylvania, 1978. The author is grateful to Professor Duffin for discussions relating to the work reported here.

[2] The author wishes to thank R. J. Duffin for reading an earlier version of this paper and making numerous suggestions for improving it, which are incorporated here. Our exposition and proofs have profited from comments of C. E. Blair and J. Borwein.

[3] Professor, College of Industrial Management, Georgia Institute of Technology, Atlanta, Georgia.

0022-3239/81/0400-0479$03.00/0

with possibly infinitely many constraints. We show, under a weak constraint qualification (see below), which is equivalent to simply the *feasibility* of (1) if $K \subseteq R^n$ is closed and all f_h for $h \in \{0\} \cup H$ are closed, that a small modification of the ordinary Lagrangian always closes the duality gap. The hypotheses that we require are uniformly weaker than the existence of a Slater point. To be more specific (see Section 5 below), we show, under such hypotheses, that there is a vector $w \in R^n$, such that

$$\lim_{\theta \searrow 0^+} \sup_{\Lambda} \inf_{x \in K} \left\{ f_0(x) + \theta w x + \sum_{h \in H} \lambda_h f_h(x) \right\} = v(P), \tag{2}$$

where $v(P)$ is the value of (1). The summation in (2) is never problematic, since only finitely many λ_h are nonzero; also, Λ is the space of all nonnegative finitely nonzero vectors $(\lambda_h | h \in H)$. The limit in (2) is a sequence limit in the reals, i.e., not a set limit. The relationship of (2) to Lagrangian-type results will be discussed in Section 5.

Thus, (2) places many duality gaps in a simple perspective: for closed, convex optimization, duality gaps can be avoided by perturbing the criterion function and sending the perturbation to zero along a ray to the origin, while leaving the constraints unchanged. For nonclosed convex optimization, some *constraint qualification* (beyond mere consistency) is necessary, but it is uniformly weaker than those necessary for a Kuhn–Tucker vector to exist.

Our methods of proof can be succinctly described, and were developed jointly in Ref. 1. We reason as follows. Since closed convex sets and since the epigraphs of closed, convex functions are describable by infinitely many linear inequalities [in the terminology of Charnes, Cooper, and Kortanek (Ref. 2), they are describable as a *semi-infinite constraint set*], this convex optimization ought to be reducible, in principle, to the study of the implied linear inequalities (*cutting-planes*) of semi-infinite systems.

Recently, Duffin and the author (Ref. 1) found methods of reducing convex programs, under a constraint qualification, to semi-infinite programs, of applying the "appropriate" result on semi-infinite systems, and then reinterpreting the resulting conclusion (which is a conclusion about the implied linear inequalities of the semi-infinite program) as a conclusion about the convex program.

This paper is very similar to Ref. 1, except that a different result about semi-infinite programs is first established here in Section 3, then applied, and then a different conclusion about the Lagrangian is obtained. Moreover, a quite closely related *limiting Lagrangian result* is due to Duffin in Ref. 3, of which this limiting Lagrangian is a refinement, in that a limiting process is taken along a line, rather than from all directions in space, and sets $K \subsetneqq R^n$ are treated.

For related work, see Blair's generalization (Ref. 4, Theorem 3) of a result from an early draft of Ref. 1, which we quoted to him, as well as McLinden's further generalization of this result in Ref. 5 to certain infinite-dimensional spaces. McLinden's work uses the elegant theory of conjugates of convex functions, as developed by Rockafellar in Ref. 6.

In the course of proving (2), we shall establish somewhat stronger results concerning the rate of convergence to the limit of (2), affine minorants which are uniformly below the functions f_h on their entire individual domains of definition, etc.

2. A Constraint Qualification Weaker than the Existence of a Kuhn–Tucker Vector

The convex program studied in this paper is (1), where H is an index set of arbitrary cardinality, K is a nonempty convex set, f_h for $h \in \{0\} \cup H$ maps a convex set $D_h \supseteq K$ into R, and D_h is the domain of f_h. In the terminology of Ref. 6, $D_h = \mathrm{dom}(f_h)$; and $f_h(x) = +\infty$, for $x \notin D_h$, would be assumed in Ref. 6. All functions f_h, $h \in \{0\} \cup H$, are convex.

Let rel int(S) denote the relative interior of the set S (Ref. 6).

We define the closure of the convex program (1) to be

$$\text{(P}'\text{)} \qquad \inf \mathrm{cl}(f_0)(x),$$

$$\text{subject to} \quad \mathrm{cl}(f_h)(x) \leq 0, \qquad \text{for all } h \in H, \tag{3}$$

$$x \in \mathrm{cl}(K),$$

where cl(S) is the closure of the set $S \subseteq R^n$, and cl(f) is the closure of the convex function f in the sense of Ref. 6. The value of (3) is denoted $v(P')$.

Here is the hypothesis that we use:

$$\text{program (1) is consistent, and } v(P) = v(P'). \tag{4}$$

Now, (4) is an exceptionally weak requirement. If, for example, all the functions f_h for $h \in \{0\} \cup H$ are closed and $K \subseteq R^n$ is closed, then (1) and (3) are the same program, and (4) simply requires that (1) be consistent. To compare (4) with the usual Slater point hypothesis, which is sufficient for a Kuhn–Tucker vector to exist when H has finite cardinality, we cite a result from Ref. 7.

Theorem 2.1. (*Ref. 7*). If there is a point $x^0 \in K^n$ satisfying:

$$f_h(x^0) \leq 0, \qquad \text{for all } h \in H,\ x^0 \in \text{rel int}(K),\ x^0 \in \text{rel int}(D_h),$$

$$\text{when } f_h \text{ is not closed, } h \in \{0\} \cup H, \tag{5}$$

then

$$v(P) = v(P').$$

Furthermore, if K is closed, then the condition that $x \in \operatorname{rel\,int}(K)$ can be dropped in (5).

Since the term

$$f_0(x) + \theta w x + \sum_{h \in H} \lambda_h f_h(x),$$

which appears in (2), need only be evaluated if $x \in K$, one makes best use of Theorem 2.1 if one takes all $D_h = K$, i.e., if one truncates all $D_h \supseteq K$ to K. Then, the condition that $x^0 \in \operatorname{rel\,int}(D_h)$ can be dropped. We retain the possibility that $D_h \supsetneqq K$ in order to obtain results stronger than (2), as mentioned in the introduction.

Since the constraint set of (3) includes that of (1), we always have

$$v(P') \leq v(P).$$

The possibility that

$$v(P') < v(P)$$

can actually occur when (5) is dropped, as shown by an example in Ref. 7 involving nonclosed convex functions. One easily verifies that the cited example does not satisfy (2) either. Thus, while (2) holds in very broad generality, it is not universally valid.

We now show that our hypothesis (4) does not imply the existence of a Kuhn–Tucker vector; in fact, it does not even imply that

$$\sup_{\Lambda} \inf_{x \in K} \left\{ f_0(x) + \sum_{h \in H} \lambda_h f_h(x) \right\} = v(P), \tag{6}$$

when

$$K = R^n, \qquad n = 2, \qquad |H| = 1.$$

Consider the convex program

$$\begin{aligned} &\inf(-y), \\ &\text{subject to} \quad (x^2 + y^2)^{1/2} - x \leq 0, \end{aligned} \tag{7}$$

which is well known as not possessing any Kuhn–Tucker vector; i.e., letting $v(P)$ denote the value of the primal problem (6), we have here that

$$v(P) = 0,$$

since the constraints have solutions

$$(x, y) = (x, 0), \qquad \text{for } x \geq 0;$$

yet there is no scalar $\lambda_1 \geq 0$ with

$$\inf_{x,y\in R} -y+\lambda_1[(x^2+y^2)^{1/2}-x] \geq v(P)=0. \tag{8}$$

In fact, for any $\lambda_1>0$, and for any specific $y_0>0$, by choosing

$$x_0=(\lambda_1 y_0^2 - 1/\lambda_1)/2,$$

we have

$$-y_0+\lambda_1((x_0^2+y_0^2)^{1/2}-x_0)=-y_0+1,$$

and thus the infimum in (8) is $-\infty$. For $\lambda_1=0$, again this infimum is $-\infty$.

This example also shows how the limiting Lagrangian [Eq. (2)] is concerned with quite a different phenomenon than is treated by the ordinary Lagrangian. The ordinary Lagrangian is related precisely to linear affine supports (or ϵ-supports) to the value function of (CP) at the origin, and the augmented Lagrangians treat the case of more general (e.g., negative definite) supporting surfaces of the perturbation function. However, in our example (7), we have, as the one-dimensional perturbation function $p(u)$,

$$p(u)=\inf\{-y|(x^2+y^2)^{1/2}-x\leq u\}=\begin{cases}+\infty \text{ (inconsistency)}, & \text{if } u<0,\\ 0, & \text{if } u=0,\\ -\infty, & \text{if } u>0.\end{cases}$$

There can be no kind of supporting surface for such a perturbation function.

The limiting Lagrangian (2) is therefore concerned with a kind of phenomenon in convex optimization which was unknown until Duffin's paper (Ref. 3), although it can be obtained from bi-function results (Ref. 6).

3. Strengthening of a Result of Blair

In this section, we strengthen Ref. 8, Corollary 2, to a form which we shall need in order to determine the implied linear inequalities of a semi-infinite system.

Let cone (S) [respectively, cl cone(S)] denote the cone [respectively, the closure of the cone] spanned by S (see Ref. 6). The following result is well known (see, e.g., Ref. 6) and is a direct application of the separating hyperplane theorem.

Lemma 3.1. For $I\neq\varnothing$ an arbitrary index set, indexing a set of vectors $\{a^i|i\in I\}$ in R^n, suppose that

$$a^i x\geq 0, \text{ all } i\in I, \qquad \text{implies } cx\geq 0, \text{ for any } x\in R^n. \tag{9}$$

Then,

$$c \in \operatorname{cl} \operatorname{cone}(\{a^i | i \in I\}).$$

Lemma 3.2. For $I \neq \varnothing$ an arbitrary index set, suppose that

$$a^i x \geq 0, \text{ all } i \in I, \qquad \text{implies } cx \geq 0. \tag{10}$$

Then, there is a vector w with the following property: For any θ, $0 < \theta \leq 1$, there is a set of nonnegative multipliers $\{\lambda_i | i \in I\}$, only finitely nonzero, such that

$$c + \theta w = \sum_{i \in I} \lambda_i a^i. \tag{11}$$

In fact, if v is any point in the relative interior of the set

$$C' = \operatorname{cone}(\{a^i | i \in I\}), \tag{12}$$

we may set

$$w = v - c. \tag{13}$$

Proof. By Lemma 3.1, $c \in \operatorname{cl} C'$; and, since v is in the relative interior of C', then $0 < \theta \leq 1$ implies that $\theta v + (1 - \theta)c$ is in the relative interior of C', by the accessibility lemma (Ref. 6); hence, it can be expressed in the form of the right-hand side of (11), with $\{\lambda_i | i \in I\}$ a finitely nonzero set of multipliers. However,

$$\theta v + (1 - \theta)c = c + \theta(v - c) = c + \theta w,$$

and so (11) holds. Since any convex set C' has a relative interior, at least one such w given by (13) exists. □

We now give our strengthening of Ref. 8, Corollary 2, which is closely related to Kortanek's *perfect duality results* (Ref. 9). Our strengthening can easily be used to prove the main result of Ref. 8, but we omit the details.

Theorem 3.1. Let $I \neq \varnothing$ be an arbitrary index set, and suppose that the system

$$a^i x \geq b_i, \qquad \text{all } i \in I, \tag{14}$$

has a solution in R^n. Suppose also that (14) implies that

$$cx \geq d, \qquad \text{for any } x \in R^n. \tag{15}$$

Then, there is a vector $w \in R^n$ and a scalar $w_0 \in R$, with the following property: For every $0 < \theta \leq 1$, there are nonnegative scalars $\{\lambda_i | i \in I\}$, only finitely nonzero, which satisfy

$$c+\theta w=\sum_{i\in I}\lambda_i a^i, \tag{16}$$

$$d+\theta w_0\leq \sum_{i\in I}\lambda_i b_i. \tag{17}$$

In fact, if $(v, -v_0)$ is any point in the relative interior of the set

$$C''=\operatorname{cone}(\{(a^i, -b_i)|i\in I\}\cup\{(0, 1)\}), \tag{18}$$

we may set

$$(w, -w_0)=(v, -v_0)-(c, -d), \tag{19}$$

i.e.,

$$w=v-c, \qquad w_0=v_0-d.$$

Proof. Since (14) is consistent, and since (14) implies (15), one easily proves that

$$a^i x\geq 0, \text{ all } i\in I, \qquad \text{implies } cx\geq 0.$$

Therefore, for $(x, r)\in R^{n+1}$ $(x\in R^n)$ arbitrary, we have that

$$r\geqslant 0,\ a^i x-b_i r\geq 0, \text{ for all } i\in I, \qquad \text{implies } cx-dr\geq 0.$$

We apply Lemma 3.2 to reach the conclusion that there exists

$$(w, -w_0)\in R^{n+1}, \qquad w\in R^n,$$

with the following property: For any $0<\theta\leq 1$, there are nonnegative scalars $\{\lambda_i|i\in I\}$, finitely nonzero, and a scalar $\varphi\geq 0$, such that

$$(c, -d)+\theta(w, -w_0)=\varphi(0, 1)+\sum_{i\in I}\lambda_i(a^i, -b_i). \tag{20}$$

Also, if $(v, -v_0)$ is any point in the relative interior of C'', we may use (19). Now, analyzing (20) by components gives (16) and (17). □

Quite clearly, if the cone C'' of (18) is closed, one may choose $w=0$ and $w_0=0$ in (16) and (17). A sufficient condition for the closure of C'' was given by Duffin and Karlovitz in Ref. 10.

4. Main Result

Our main result (Theorem 4.1 below) is obtained by applying Theorem 3.3 to a *semi-infinite system* of linear inequalities equivalent to (3), and then interpreting this outcome by methods of algebraic manipulation developed in Ref. 1. We reproduce the latter here, for the sake of a self-contained presentation.

Under the hypothesis (4), we may assume that K is closed, as are the functions f_h for $h \in \{0\} \cup H$; i.e., we may deal with (3) in place of (1). This is assumed throughout the remainder. Therefore, we have representations via hyperplanes:

$$K = \{x \in R^n | a^j x \geq a_0^j, j \in I(-1)\}, \tag{21}$$

$$\operatorname{epi}(f_j) = \{(z, x) \in R^{n+1}, x \in R^n | b_0^j z + a^j x \geq a_0^j, j \in I(h)\}, \tag{22}$$

for index sets

$$I(h), h \in \{-1\} \cup \{0\} \cup H,$$

where possibly

$$I(-1) = \varnothing,$$

but

$$I_h \neq \varnothing, \qquad \text{for } h \neq -1,$$

and as usual $\operatorname{epi}(f_h)$ denotes the epigraph of f_h:

$$\operatorname{epi}(f_h) = \{(z, x) \in R^{n+1}, x \in R^n | z \geq f_h(x)\}, \tag{23}$$

which is a closed, convex set. Obviously, in (22),

$$b_0^j \geq 0, \qquad \text{for all } j \in I(h) \text{ and } h \in \{0\} \cup H.$$

Lemma 4.1. (*Ref. 1*). Fix $h \in H$, and suppose that $\operatorname{epi}(f_h)$ is given by (22). Then, for any $x \in D_h$, $f_h(x) \leq 0$ is equivalent to the semi-infinite system

$$a^j x \geq a_0^j, \qquad j \in I(h). \tag{24}$$

Proof. We have

$$\begin{aligned} f_h(x) \leq 0 &\leftrightarrow (0, x) \in \operatorname{epi}(f_h) \\ &\leftrightarrow b_0^j \cdot 0 + a_j x \geq a_0^j, \qquad \text{all } j \in I(h) \\ &\leftrightarrow a^j x \geq a_0^j, \qquad \text{all } j \in I(h). \end{aligned}$$

□

Corollary 4.1. Assume that the condition (4) holds. Then, the value $v(P)$ of the convex program is also the value of this semi-infinite program:

$$\begin{aligned} &\inf z, \\ &\text{subject to} \quad b_0^j z + a^j x \geq a_0^j, \qquad j \in I(0) \\ &\qquad\qquad a^j x \geq a_0^j, \qquad \text{for all } j \in I(h) \\ &\qquad\qquad\qquad\qquad \text{and } h \in \{-1\} \cup H. \end{aligned} \tag{25}$$

Proof. It is immediate from Lemma 4.1. □

It is now the point to complete the program outlined at the beginning of this paper and invoke Theorem 3.3.

Theorem 4.1. Suppose that the condition (4) holds and $v(P)$ is finite. Then, there exists w_0, $w_1 \in R$ and $w \in R^n$, with the following property: For any scalar θ in the range $0 < \theta \leq 1$, there exist $\gamma \in R^n$, $\gamma^0 \in R$, and nonnegative scalars $\{\lambda_h \mid h \in H\}$, only finitely many of which are nonzero, and $\beta^h \in R^n$, $\beta_0^h \in R$ for $h \in \{0\} \cup H$, satisfying the three conditions below:

Condition C1. $\gamma x + \gamma_0 \leq 0$, for $x \in K$.

Condition C2. $\beta^h x + \beta_0^h \leq f_h(x)$, for all $x \in D_h$ and $h \in \{0\} \cup H$.

Condition C3. $\gamma x + \gamma^0 + (1 + \theta w_0)(\beta^0 x + \beta_0^0) + \theta(wx + w_1)$

$$+ \sum_{h \in H} \lambda_h(\beta^h x + \beta_0^h) \geq v(P), \qquad \text{for all } x \in R^n.$$

In fact, w_0, w_1, w can be chosen arbitrarily to satisfy

$$(w_0, w, w_1) = (v_0, v, v_1) - (1, 0, -v(P)), \tag{26}$$

where (v_0, v, v_1), with v_0, $v_2 \in R$ and $v \in R^n$, is any point in the relative interior of

$$\begin{aligned} C = \operatorname{cone}(\{(b_0^j, a^j, -a_0^j) \mid j \in I(0)\} \cup \bigcup_{h \in \{-1\} \cup H} \{(0, a^j, -a_0^j) \mid j \in I(h)\} \\ \cup \{(0, 0, 1)\}). \end{aligned} \tag{27}$$

Proof. By Corollary 4.1 and by Theorem 3.1, (16) and (17), which we saw is equivalent to (20), we have, for $0 < \theta \leq 1$,

$$\begin{aligned} (1, 0, -v(P)) + \theta(w_0, w, w_1) = \varphi(0, 0, 1) + \sum_{j \in I(0)} \varphi_j(b_0^j, a^j, -a_0^j) \\ + \sum_{h \in \{-1\} \cup H} \sum_{j \in I(h)} \varphi_j(0, a^j, -a_0^j), \end{aligned} \tag{28}$$

since (25) implies

$$z \cdot 1 + x \cdot 0 \geq v(P).$$

Of course, in (28), $\varphi \geq 0$ and all $\varphi_j \geq 0$, and only finitely many of the quantities φ_j for all $j \in I(h)$ and all $h \in \{-1, 0\} \cup H$ are actually nonzero. Also, (w_0, w, w_1) is any solution to (26), such that (v_0, v, v_1) is in the relative interior of C of (27).

We now analyze (28) by the methods of Ref. 1. From the first components in (28), we obtain

$$1 + \theta w_0 = \sum_{\substack{j \in I(0) \\ b_0^j > 0}} \varphi_j b_0^j. \tag{29}$$

We will, in general, define for $h \in \{0\} \cup H$

$$\lambda_h = \sum_{\substack{j \in I(h) \\ b_0^j > 0}} \varphi_j b_0^j, \tag{30}$$

with the understanding that

$$\lambda_h = 0, \qquad \text{if } b_0^j = 0, \text{ for all } j \in I(h), \text{ with } \varphi_j > 0.$$

Thus, we have

$$\lambda_0 = 1 + \theta w_0,$$

from (29). Clearly, only finitely many λ_h can be nonzero, and all $\lambda_h \geq 0$, by the conditions on the scalars $\varphi_j \geq 0$.

Next, we define these vectors and scalars:

$$\gamma = -\sum_{j \in I(-1)} \varphi_j a^j - \sum_{h \in \{0\} \cup H} \sum_{\substack{j \in I(h) \\ b_0^j = 0}} \varphi_j a^j, \tag{31}$$

$$\gamma_0 = \sum_{j \in I(-1)} \varphi_j a_0^j + \sum_{h \in \{0\} \cup H} \sum_{\substack{j \in I(h) \\ b_0^j = 0}} \varphi_j a_0^j, \tag{32}$$

where an empty summation is zero. From (21), if $j \in I(-1)$,

$$a^j x \geq a_0^j, \qquad \text{for } x \in K;$$

and from (22), if $b_0^j = 0$, we have again

$$a^j x \geq a_0^j, \qquad \text{for } j \in I(h) \text{ and } h \in \{0\} \cup H,$$

as $x \in D_h \supseteq K$. Hence,

$$\gamma x + \gamma_0 \leq 0, \qquad \text{for } x \in K;$$

i.e., Condition C1 holds.

For $h \in \{0\} \cup H$, if $b_0^j > 0$, since

$$b_0^j f_h(x) + a^j x \geq a_0^j, \qquad \text{for } x \in D_h,$$

we have

$$f_h(x) \geq (a^j/b_0^j)x + a_0^j/b_0^j, \qquad \text{for } x \in D_h. \tag{33}$$

Combining (30) and (33) for $\lambda_h > 0$, we have

$$\begin{aligned} \lambda_h f_h(x) &= \sum_{\substack{j \in I(h) \\ b_0^j > 0}} \varphi_j b_0^j f_h(x) \\ &\geq \sum_{\substack{j \in I(h) \\ b_0^j > 0}} \varphi_j b_0^j(-(a^j/b_0^j)x + a_0^j/b_0^j), \qquad \text{for } x \in D_h, \end{aligned} \tag{34}$$

and so defining

$$\beta^h x + \beta_0^h = (1/\lambda_h) \sum_{\substack{j \in I(h) \\ b_0^j > 0}} \varphi_j b_0^j(-(a^j/b_0^j)x + a_0^j/b_0^j), \tag{35}$$

we obtain Condition C2. To be precise, we actually have

$$\mathrm{cl}(f_h)(x) \geq \beta^h x + \beta_0^h, \qquad \text{for all } x \in D_h;$$

but, since

$$\mathrm{cl}(f_h)(x) \leq f_h(x), \qquad \text{for all } x \in D_h,$$

Condition C2 follows.

If $\lambda_h = 0$, one can arbitrarily pick $\beta^h x + \beta_0^h$ to satisfy Condition C2, at least one such affine form existing since f_h is somewhere finite, and hence has at least one subgradient at one point.

For the concluding part of our analysis, we write (28) again with the first component dropped, in this form:

$$\begin{aligned} (0, -v(P)) + \theta(w, w_1) &= \varphi(0, 1) + (-\gamma, -\gamma_0) \\ &\quad + \sum_{h \in \{0\} \cup H} \sum_{\substack{j \in I(h) \\ b_0^j > 0}} \varphi_j b_0^j (a^j/b_0^j, -a_0^j/b_0^j) \\ &= \varphi(0, 1) + (-\gamma, -\gamma_0) + \sum_{h \in \{0\} \cup H} \lambda_h(-\beta^h, -\beta_0^h). \end{aligned} \tag{36}$$

It remains only to dot product both sides of (36) with $(-x, -1)$, where $x \in R^n$ is arbitrary, to obtain

$$v(P) + \theta(-wx - w_1) = -\varphi + (\gamma x + \gamma_0) + \sum_{h \in \{0\} \cup H} \lambda_h(\beta^h x + \beta_0^h), \tag{37}$$

for all $x \in R^n$. Since $\varphi \geq 0$, (37) immediately yields Condition C3, using

$$\lambda_0 = 1 + \theta w_0. \qquad \square$$

Corollary 4.2. Suppose that the condition (4) holds for the program (1) and that $v(P)$ is finite. Then, there exists w_0, $w_1 \in R$ and $w \in R^n$, with the following property: For any scalar in the range $0 < \theta \leq 1$, there exist nonnegative scalars $\{\lambda_h \mid h \in H\}$, only finitely many of which are nonzero, with

$$(1 + \theta w_0) f_0(x) + \theta(wx + w_1) + \sum_{h \in H} \lambda_h f_h(x) \geq v(P), \qquad \text{for all } x \in K. \tag{38}$$

Furthermore, w_0, w_1, w can be arbitrarily chosen, subject to the condition (26), where (v_0, v, v_1) is a point in the relative interior of the convex set C of (27).

Proof. It is immediate from Theorem 4.1. $\square$

5. Limiting Lagrangian Equation (2)

In the usual Lagrangian and its associated Kuhn–Tucker theory, typically one seeks sufficient conditions for the equality

$$\max_{\substack{\lambda_h \geq 0 \\ h \in H}} \inf_{x \in K} \left\{ f_0(x) + \sum_{h \in H} \lambda_h f_h(x) \right\} = v(P), \tag{39}$$

where, in some instances, the *max* (maximum) is relaxed to a *sup* (supremum). The usual theory requires the cardinality of H to be finite.

What our Corollary 4.2 is concerned with are relations more complex than (39), due to the presence of one more operation on the left-hand side: the taking of a limit. We next establish an inequality which will help us make this point.

Lemma 5.1. Suppose that $w(\theta) \in R^n$ is defined for $0 < \theta \leq 1$, and the set of all vectors of this form is bounded. Then, if (CP) is consistent and has finite value $v(P)$,

$$\limsup_{\theta \searrow 0^+} \sup_{\Lambda} \inf_{x \in K} \{ f_0(x) + \theta w(\theta) x + \sum_{h \in H} \lambda_h f_h(x) \} \leq v(P). \tag{40}$$

Proof. First, observe that, for any θ and element of Λ,

$$\inf_{x \in K} \left\{ f_0(x) + \theta w(\theta) x + \sum_{h \in H} \lambda_h f_h(x) \right\} \leq \inf_{x \in K} \{ f_0(x) + \theta w(\theta) x \mid f_h(x) \leq 0,\ h \in H \}, \tag{41}$$

since all $\lambda_h \geq 0$, if $(\lambda_h \mid h \in H) \in \Lambda$. Therefore, the left-hand side in (40) does not exceed

$$\limsup_{\theta \searrow 0^+} \inf_{x \in K} \{f_0(x) + \theta w(\theta)x \mid f_h(x) \leq 0,\ h \in H\}. \tag{42}$$

We first consider the case that $v(P)$ is finite.

Next, let $x^{(n)}$ be chosen so that

$$f_h(x^{(n)}) \leq 0, \qquad \text{for } h \in H,$$

and

$$f_0(x^{(n)}) \leq v(P) + 1/n,$$

which is possible, since $v(P)$ is the value of (1). We see that (42) does not exceed

$$\limsup_{\theta \searrow 0^+} \{f_0(x^{(n)}) + \theta w(\theta) x^{(n)}\} = f_0(x^{(n)}) \leq v(P) + 1/n, \tag{43}$$

using the boundedness condition of $w(\theta)$. Now if $v(P) + 1/n$ is an upper bound on the left-hand side of (43) for any n, so is $v(P)$. This establishes (40). If $v(P) = -\infty$, a quite similar argument obtains this result. □

From (38) of Corollary 4.2, for any θ in the range $0 < \theta \leq 1$, such that

$$1 + \theta w_0 > 0,$$

we have, upon division by $1 + \theta w_0$,

$$\inf_{x \in K} \left\{ f_0(x) + \theta' w x + \sum_{h \in H} \lambda'_h f_h(x) \right\} \geq v(P)/(1 + \theta w_0) - \theta' w_1, \tag{44}$$

for suitable $\lambda'_h \geq 0$, where we have set

$$\theta' = \theta/(1 + \theta w_0), \qquad \lambda'_h = \lambda_h/(1 + \theta w_0),$$

in terms of the quantities of (38). Therefore,

$$\liminf_{\theta' \searrow 0^+} \sup_{\Lambda} \inf_{x \in K} \left\{ f_0(x) + \theta' w x + \sum_{h \in H} \lambda'_h f_h(x) \right\} \geq v(P). \tag{45}$$

Putting together (40) and (45) above, we obtain (under the hypotheses of Corollary 4.4), after renaming θ' to θ and λ'_h to λ_h, Eq. (2). Comparing the standard Lagrangian result (39) with ours (2), we see their similarity in nature. The limit appearing in (2) suggests the term limiting Lagrangian equation for (2).

Note that, if $K \subseteq R^n$ is bounded, (2) becomes the ordinary Lagrangian statement (6), since $\theta w x \to 0$ uniformly in $x \in K$ as $\theta \to 0$.

To put these results in perspective, a main feature of the ordinary Lagrangian is that it "reduces" a constrained optimization to an unconstrained one. The limiting Lagrangian reduces a constrained optimization to a sequence of unconstrained ones, as $\theta \downarrow 0^+$. The ordinary Lagrangian, when a Kuhn–Tucker vector exists, allows an economic interpretation of the dual: with the correct "prices" of factors, the constraints can be dropped. Part of the economic interpretation of the limiting Lagrangian is that, with the correct prices *and* perturbation of the criterion function, one obtains an unconstrained profit maximization which is *almost* the same (in value) as the constrained one.

As regards the determination of $w \in R^n$ of (2), we first discuss a few points concerning the determination of relative interior points (v_0, v, v_1) of C in (27). We do not try here to give an efficient algorithm; we merely wish to indicate that these quantities are often, in principle, computable.

For vectors $v^i \in R^q$ and a nonempty index set $I \neq \emptyset$, the determination of a relative interior point of

$$C_I = \text{cone}(\{v^i \mid i \in I\}) \tag{46}$$

is never, in principle, problematic, once one has some spanning set, say $\{v^1, \ldots, v^t\}$ of $\{v^i \mid i \in I\}$, in the sense of a vector space span. A relative interior point of C_I of (51) is always given by

$$v = v^1 + \cdots + v^t. \tag{47}$$

Indeed, for any vector

$$w = \sum_{i=1}^{t} \rho_i v^i$$

in the vector space spanned by $\{v^1, \ldots, v^t\}$, i.e., in the manifold spanned by C_I of (46), there is a sufficiently small $\epsilon > 0$, so that

$$1 + \epsilon\rho_i > 0, \qquad \text{for } i = 1, \ldots, t, \tag{48}$$

and hence

$$v + \epsilon w \in C_I.$$

Since w in the manifold spanned by C_I is arbitrary, v of (47) is a relative interior point of C_I, by standard criteria (see, e.g., Ref. 6).

To obtain $\{v^1, \ldots, v^t\}$, it often suffices to know t, the dimension of the manifold spanned by C_I. Here, we have in mind primarily the case that a countable dense subset $\{v^i \mid i \in I'\}$, which can be effectively listed, can be effectively extracted from $\{v^i \mid i \in I\}$. Since

$$\text{cl cone}(\{v^i \mid i \in I'\}) \supseteq \text{cone}(\{v^i \mid i \in I\}),$$

there are also t linearly independent vectors in $\{v^i \mid i \in I'\}$. Thus, one can simply continue a listing until t independent ones are found. We avoid details on the points raised in this paragraph, since a full discussion of these matters requires a knowledge of recursion theory, which we do not assume here.

The simplest case in $t = q$; i.e., the cone C_I of (51) is fully dimensional, and our next result shows that this is indeed a very common case for the cone C of (27). Before we begin the proof of our next result, one may remark that full-dimensionality occurs if

$$v^i x = 0, \text{ for all } i \in I, \qquad \text{implies } x = 0. \tag{49}$$

Indeed, if (49) holds, but the linear span of $\{v^i \mid i \in I\}$ is a subspace $L \subsetneq R^q$, its perpendicular subspace $L^\perp$ has a nonzero vector $\bar{x}$. Then,

$$v^i \bar{x} = 0, \qquad \text{for all } i \in I,$$

but $\bar{x} \neq 0$, a contradiction.

Proposition 5.1. Suppose that (1) is feasible and has finite value $v(P)$. Barring the case that there exists a nonzero vector x^* such that, for all $\theta \in R$, we have, for any solution $\bar{x}$ to (CP),

$$f_h(\bar{x} + \theta x^*) \leq 0, \qquad \text{all } h \in H, \tag{50}$$

$$f_0(\bar{x} + \theta x^*) = f_0(\bar{x}), \tag{51}$$

$$\bar{x} + \theta x^* \in K, \tag{52}$$

then the cone C of (27) is of full dimension $n + 2$.

Proof. We have to show that there is no nonzero solution to all the equalities

$$\begin{aligned} b_0^j z + a^j x - a_0^j w &= 0, \qquad j \in I(0), \\ a^j x - a_0^j w &= 0, \qquad j \in I(h) \text{ and } h \in \{-1\} \cup H, \\ w &= 0, \end{aligned} \tag{53}$$

or, equivalently, to the equalities

$$\begin{aligned} b_0^j z + a^j x &= 0, \qquad j \in I(0), \\ a^j x &= 0, \qquad j \in I(h) \text{ and } h \in \{-1\} \cup H. \end{aligned} \tag{54}$$

Since at least one $b_0^j > 0$, for $j \in I(0)$, as f_0 has a subgradient at least one point, we cannot have $x = 0$ in a nonzero solution to (54).

Suppose that (z^*, x^*) solves (59), so that $x^* \neq 0$. By homogeneity, we may assume that $z^* \leq 0$. Let $\bar{x}$ be any solution to (1). Then, for any $\theta \geq 0$, fixing $h \in \{-1\} \cup H$ and letting $j \in I(h)$ be arbitrary, we have

$$a^j(\bar{x} + \theta x^*) = a^j\bar{x} + \theta a^j x^* = a^j\bar{x} \geq a_0^j, \tag{55}$$

since

$$a^j x^* = 0,$$

from (54), and

$$a^j\bar{x} \geq a_0^j,$$

by Lemma 4.1. Thus, by Lemma 4.1, we have (50) and (52) for $\theta \geq 0$. Also, since

$$b_0^j f(\bar{x}) + a^j\bar{x} \geq a_0^j, \qquad \text{for any } j \in I(0),$$

as

$$(f_0(\bar{x}), \bar{x}) \in \operatorname{epi}(f_0)),$$

we have

$$\begin{aligned} b^j(f_0(\bar{x}) + \theta z^*) + a^j(\bar{x} + \theta x^*) &= (b^j f_0(\bar{x}) + a^j\bar{x}) + \theta(b^j z^* + a^j x^*) \\ &= b^j f_0(\bar{x}) + a^j\bar{x} \geq a_0^j, \end{aligned} \tag{56}$$

since

$$b^j z^* + a^j x^* = 0.$$

This gives (51) for $\theta \geq 0$, with = replaced by $\leq$, since

$$(f_0(\bar{x}) + \theta z^*, \bar{x} + \theta x^*) \in \operatorname{epi}(f_0)$$

and

$$z^* \leq 0, f_0(\bar{x}) + \theta z^* \leq f(\bar{x}), \qquad \text{for all } \theta \geq 0.$$

Now, if $z^* < 0$, from the above, $f_0(\bar{x} + \theta x^*)$ can be indefinitely decreased by sending $\theta \nearrow +\infty$, and all the while $\bar{x} + \theta x^*$ is feasible in (1). This contradicts that (1) has finite value. Hence, $z^* = 0$, and we can repeat the analysis with $(-z^*, -x^*)$ replacing (z^*, x^*), and in this manner obtain (50) and (52) for all $\theta \in R$. We also obtain

$$f_0(\bar{x} + \theta x^*) \leq f_0(\bar{x}), \qquad \text{for all } \theta \in R;$$

since f_0 is convex, we clearly have (51). □

Thus, if it is known that the feasible region contains no full line, or that

$$|f(x)| \nearrow +\infty, \qquad \text{as } \|x\| \nearrow +\infty,$$

or that f_0 is not constant on any line, or that f_0 is not constant on any line in the feasible region of (1), all of these being commonly occurring hypotheses, Proposition 5.3 shows that the dimension of the cone C of (27) is full, i.e., is $n+2$.

From (26), once an interior point (v_0, v, v_1) is found, we can compute w for use in the limiting Lagrangian as $w = v$.

If the value of the dual is $+\infty$, then we prove that (1) is not feasible quite easily. Indeed, if x^* is a feasible solution to (1), we have, under a bounded ness assumption for $w(\theta)$, $0<\theta\leq 1$,

$$\begin{aligned} &\lim_{\theta\searrow 0^+} \sup_{\Lambda} \inf_{x\in K}\left\{ f_0(x)+\theta w(\theta)x + \sum_{h\in H} \lambda_h f_h(x)\right\} \\ &\quad\leq \lim_{\theta\searrow 0^+} \sup_{\Lambda}\left\{ f_0(x^*)+\theta x(\theta)x^* + \sum_{h\in H} \lambda_h f_h(x^*)\right\} \\ &\quad\leq \lim_{\theta\searrow 0^+} \{f_0(x^*)+\theta w(\theta)x^*\} = f_0(x^*). \end{aligned} \tag{57}$$

This would contradict a value of $+\infty$ in the dual.

Of course, as in any duality theory, there is the possibility that both primal and dual are inconsistant, i.e., the primal is inconsistent (nominally given value $+\infty$) and the dual has value $-\infty$.

References

1. DUFFIN, R. J., and JEROSLOW, R. G., *Affine Minorants and Lagrangian Functions*, Mathematical Programming (to appear).
2. CHARNES, A., COOPER, W. W., and KORTANEK, K. O., *On Representations of Semi-Infinite Programs Which Have No Duality Gaps*, Management Science, Vol. 12, No. 1, 1965.
3. DUFFIN, R. J., *Convex Analysis Treated by Linear Programming*, Mathematical Programming, Vol. 4, No. 2, 1973.
4. BLAIR, C. E., *Convex Optimization and Lagrange Multipliers*, Mathematical Programming, Vol. 15, pp. 87–91, 1978.
5. MCLINDEN, L., *Affine Minorants Minimizing the Sum of Convex Functions*, Journal of Optimization Theory and Applications, Vol. 24, No. 4, 1978.
6. ROCKAFELLAR, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
7. BLAIR, C. E., BORWEIN, J., and JEROSLOW, R. G., *Convex Programs and Their Closures*, 1978 (manuscript).
8. BLAIR, C. E., *A Note on Infinite Systems of Linear Inequalities in R^n*, Journal of Mathematical Analysis and Applications, Vol. 48, No. 1, 1974.
9. KORTANEK, K. O., *Constructing a Perfect Duality in Infinite Programming*, Applied Mathematics and Optimization, Vol. 3, No. 4, 1977.
10. DUFFIN, R. J., and KARLOVITZ, L. A., *An Infinite Linear Program with a Duality Gap*, Management Science, Vol. 12, No. 1, 1965.