Reprinted from JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS Vol. 7, No. 3, March 1971
Printed in Belgium

Ordinary Convex Programs Without a Duality Gap[1]

R. T. ROCKAFELLAR[2]

Communicated by M. R. Hestenes

Abstract. In the Kuhn–Tucker theory of nonlinear programming, there is a close relationship between the optimal solutions to a given minimization problem and the saddlepoints of the corresponding Lagrangian function. It is shown here that, if the constraint functions and objective function are *faithfully convex* in a certain broad sense and the problem has feasible solutions, then the *inf sup* and *sup inf* of the Lagrangian are necessarily equal.

Let C be a nonempty convex subset of R^n, and let $f_0, f_1, ..., f_m$ be real-valued, convex functions on C. The ordinary convex program

$$(P) \qquad \text{minimize } f_0(x) \text{ over } C \text{ subject to } \quad f_1(x) \leqslant 0, ..., f_m(x) \leqslant 0$$

has as its dual, in the sense of conjugate-function theory (Refs. 1–2), the problem

$$(P^*) \qquad \text{maximize } g(y) \text{ over } R_+^m,$$

where R_+^m is the nonnegative orthant of R^m, and g is the extended-real-valued, concave function on R_+^m defined by

$$g(y) = \inf\{L(x, y) \mid x \in C\}, \tag{1}$$

$$L(x, y) = f_0(x) + y_1 f_1(x) + \cdots + y_m f_m(x). \tag{2}$$

The dual problem is important in computational methods which solve (P) using Lagrange multipliers: typically, one maximizes g by some algorithm which involves repeated calculation of the infimum in (1) [see Geoffrion

[1] Paper received April 28, 1970. This work was supported in part by the Air Force Office of Scientific Research under Grant No. AF-AFOSR-1202-67B.

[2] Professor of Mathematics, University of Washington, Seattle, Washington.

(Ref. 3) for references and a general discussion]. For such methods to succeed, it is essential that there be no *duality gap*. In other words, the infimum in (P) and the supremum in (P^*) must be equal. It is therefore of interest to know under what conditions one can be sure that there is no duality gap.

Conditions of this sort have been developed by many authors. The conditions usually also entail the existence of optimal solutions to either (P) or (P^*), although this would not be required by most algorithms that seek approximate solutions to (P).

The well-known theorem of Kuhn and Tucker (Ref. 4) asserts that

$$\min(P) = \max(P^*),$$

under the assumption that (P) has an optimal solution at which the functions f_i are differentiable and satisfy a constraint qualification. Fan, Glicksberg, and Hoffman (Ref. 5) have shown much more generally that

$$\inf(P) = \max(P^*),$$

under the simple assumption that the constraints in (P) can be satisfied with strict inequality (Slater condition). This result has been extended to allow for linear equation constraints, either explicit or implicit (see Ref. 2, Section 28).

Theorems of the type

$$\min(P) = \sup(P^*) \tag{3}$$

have been developed by Rockafellar (Ref. 6) in terms of growth properties of C and the functions f_i. In particular, (3) is known to hold if C is closed, each f_i is lower semicontinuous, and there exist real numbers u_i, $i = 0, 1,..., m$, such that the convex set

$$\{x \in C \mid f_i(x) \leqslant u_i \quad \text{for} \quad i = 0, 1,..., m\} \tag{4}$$

is nonempty and bounded.

Other results about duality gaps have been obtained through the study of the perturbation function p for (P), where

$$p(u) = \inf\{f_0(x) \mid x \in C, \ \ f_1(x) \leqslant u_1 ,..., \ \ f_m(x) \leqslant u_m\}. \tag{5}$$

(see Ref. 2, Section 29.) If (P) is consistent, a necessary and sufficient condition for there to be no duality gap is that p be lower semicontinuous at $u = 0$.

The purpose of the present paper is to point out a large and important class of convex programs for which consistency alone guarantees that there is no duality gap. For problems in this class, it is unncessary to check any further assumptions concerning sets of the form (4) or to verify directly any properties of the function p.

Theorem. Suppose that $C = R^n$ and that each of the functions f_i satisfies the following regularity condition: f_i is not affine (linear-plus-a-constant) along any line segment, unless f_i is affine along the entire line extending the line segment. If (P) is consistent, then

$$\inf(P) = \sup(P^*)$$

or, in other words,

$$\inf_{x \in R^n} \sup_{y \in R_+^m} L(x, y) = \sup_{y \in R_+^m} \inf_{x \in R^n} L(x, y).$$

Observe that the regularity condition in this theorem is satisfied by f whenever f_i is linear or quadratic. In fact, it is satisfied whenever f_i is *analytic*. Thus, the theorem is applicable in particular to all convex programs on $C = R^n$ with analytic objective and analytic constraints.

Of course, the regularity condition does not actually require any differentiability at all. It is satisfied, as one can easily verify, if, and only if, every f_i can be expressed in the form

$$f_i(x) = h_i(A_i x) + l_i(x),$$

where h_i is a finite, strictly convex function on R^{n_i}, A_i is a linear transformation from R^n to R^{n_i}, and l_i is an affine function on R^n. The term $h_i(A_i x)$ may be omitted entirely, or A_i may be the identity transformation, $n_i = n$. On the other hand, $l_i(x)$ may be a constant, perhaps 0.

Proof. Let α be the infimum in (P). We can assume that α is finite, since the result is trivial otherwise. Let I_0 be the set of indices i in $\{1,..., m\}$ such that (P) has at least one feasible solution x with $f_i(x) < 0$. Let I_1 be the complement of I_0 in $\{1,..., m\}$, and let

$$C_0 = \{x \in R^n \mid f_i(x) \leqslant 0, \quad i \in I_1\}.$$

Then, (P) has a feasible solution x with $f_i(x) < 0$ for every $i \in I_0$ [that is, $x = (1/n) \sum_{i \in I_0} x_i$, where x_i is, for each $i \in I_0$, a feasible solution with $f_i(x_i) < 0$]. On the other hand, f_i is identically zero on C_0 for each $i \in I_1$. [If one had $x_0 \in C_0$ and $f_k(x_0) < 0$ for a certain $k \in I_1$, then for small $\epsilon > 0$ the point $x' = (1 - \epsilon)x + \epsilon x_0$, where x is a feasible solution with $f_i(x) < 0$ for every $i \in I_0$, would be a feasible solution with $f_k(x') < 0$, contradicting $k \notin I_0$.]

Problem (P) is equivalent to the ordinary convex program

$$(P_0) \qquad \text{minimize } f_0(x) \text{ over } C_0 \text{ subject to } f_i(x) \leqslant 0 \text{ for } i \in I_0 .$$

There is at least one feasible solution x to (P_0) with $f_i(x) < 0$ for every $i \in I_0$. Thus, (P_0) is strictly consistent, and it follows from the theorem of Fan, Glicksberg, and Hoffman that there exist Lagrange multipliers $\bar{y}_i \geqslant 0$, $i \in I_0$, such that

$$\inf_{x \in C_0} \left\{ f_0(x) + \sum_{i \in I_0} \bar{y}_i f_i(x) \right\} = \max(P_0{}^*) = \inf(P_0) = \alpha.$$

Let M be the lineality space of C_0, the subspace of R^n consisting of all the vectors z such that $C_0 + z = C_0$ (Ref. 2, Section 8). Define $\bar{f}_0$ on R^n by

$$\bar{f}_0(x) = \inf_{z \in M} f(x + z), \tag{6}$$

where

$$f = f_0 + \sum_{i \in I_0} \bar{y}_i f_i . \tag{7}$$

Then, $\bar{f}_0$ is a convex function, because f is convex (see Ref. 2, Section 8), and we have

$$\inf_{x \in C_0} \bar{f}_0(x) = \inf_{x \in C_0} f(x) = \alpha. \tag{8}$$

Note that $\bar{f}_0$ is necessarily finite everywhere, since, if not, f_0 would have to be identically $-\infty$ (Ref. 2, Theorem 7.2), contrary to the assumption that the infimum in (P) was finite. Furthermore, the definition (6) implies that

$$\bar{f}_0(x + z) = \bar{f}_0(x) \quad \text{if} \quad z \in M. \tag{9}$$

We proceed now to apply the Lagrange multiplier theorem in Ref. 6, p. 39 to the ordinary convex program

$$(P_1) \qquad \text{minimize } \bar{f}_0 \text{ over } R^n \text{ subject to } f_i(x) \leqslant 0 \text{ for } i \in I_1 .$$

The infimum in (P_1) is α by (8). To verify that the hypothesis of this theorem is satisfied, suppose that z is a recession vector (Ref. 2, Section 8) common to $\bar{f}_0$ and the functions f_i, $i \in I_1$. In other words, z possesses the property that, for every $x \in R^n$, one has

$$\bar{f}_0(x + z) \leqslant \bar{f}_0(x) \qquad \text{and} \qquad f_i(x + z) \leqslant f_i(x), \qquad i \in I_1 . \tag{10}$$

Then, in particular, the half-line $\{x + \lambda z \mid \lambda \geqslant 0\}$ is contained in C_0, the set of feasible solutions to (P_1), where x is an arbitrary element of C_0. The functions f_i, $i \in I_1$, then vanish on this half-line and, hence, by our regularity assumption, they vanish on the line extending this half-line. Thus, the expressions

$$f_i(x + \lambda z), \qquad i \in I_1,$$

are constant as functions of λ if $x \in C_0$ and, consequently, they are constant as functions of λ for every $x \in R^n$ (Ref. 2, Section 8). Therefore, $z \in M$, and $-z$ is also a recession vector common to $\bar{f}_0$ and the functions f_i, $i \in I_1$. This verifies that the hypothesis of the cited theorem is satisfied, and we may conclude that

$$\min(P_1) = \sup(P_1{}^*).$$

Thus, α is the supremum of

$$\inf_{x \in R^n} \Big\{ \bar{f}_0(x) + \sum_{i \in I_1} y_i f_i(x) \Big\} \tag{11}$$

over all choices of $y_i \geqslant 0$, $i \in I_1$. Since

$$f_i(x + z) = f_i(x) \qquad \text{if} \quad z \in M, \quad i \in I_1, \tag{12}$$

the latter supremum is the same as the supremum of the expression

$$\inf_{x \in R^n} \Big\{ f(x) + \sum_{i \in I_1} y_i f_i(x) \Big\} = \inf_{x \in R^n} \Big\{ f_0(x) + \sum_{i \in I_0} \bar{y}_i f_i(x) + \sum_{i \in I_1} y_i f_i(x) \Big\} \tag{13}$$

over all $y_i \geqslant 0$, $i \in I_1$. In other words, we have

$$\alpha = \sup_{y \in R_+^m} \inf_{x \in R^n} \Big\{ f_0(x) + \sum_{i=1}^{m} y_i f_i(x) \Big\}, \tag{14}$$

and the theorem is thereby proved.

References

1. Rockafellar, R. T., *Duality in Nonlinear Programming*, Mathematics of the Decision Sciences, Part I, Edited by G. B. Dantzig and A. F. Veinott, American Mathematical Society, Providence, Rhode Island, 1968.
2. Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.

3. GEOFFRION, A., *Duality in Nonlinear Programming: A Simplified Applications-Oriented Development*, University of California at Los Angeles, Western Management Science Institute, Working Paper No. 150, 1969.
4. KUHN, H. W., and TUCKER, A. W., *Nonlinear Programming*, Proceedings of the Second Berkeley Symposium on Mathematics, Statistics and Probability, Edited by J. Neyman, University of California Press, Berkeley, California, 1951.
5. FAN, K., GLICKSBERG, I., and HOFFMAN, A. J., *Systems of Inequalities Involving Convex Functions*, Proceedings of the American Mathematical Society, Vol. 8, No. 3, 1957.
6. ROCKAFELLAR, R. T., *Helly's Theorem and Minima of Convex Functions*, Duke Mathematical Journal, Vol. 32, No. 3, 1965.

Printed in Belgium by the St. Catherine Press Ltd., Tempelhof 37, Bruges