

A DUALITY THEOREM FOR NON-LINEAR PROGRAMMING*

BY

PHILIP WOLFE

The RAND Corporation

Summary. A dual problem is formulated for the mathematical programming problem of minimizing a convex function under convex constraints which reduces to the classical dual problem in the case of linear programming problems. Duality theorems are proved regarding the relationship between the problem and its dual.

1. Introduction. A duality theorem in mathematical programming is, generally speaking, the statement of a relationship of a certain kind between two mathematical programming problems. The relationship commonly has three aspects: (a) one problem—the “primal”—is a constrained minimization problem, and the other—the “dual”—is a constrained maximization problem; (b) the existence of a solution to one of these problems ensures the existence of a solution to the other, in which case their respective extreme values are equal; and (c) if the constraints of one problem are consistent, while those of the other are not, there is a sequence of points satisfying the constraints of the first on which its objective function tends to infinity.

Dennis [1] and Dorn [2, 3] have formulated dual problems for the primal problem of minimizing a convex function f of several real variables under linear constraints, and shown (b) above in the case that f is quadratic, or just differentiable but strictly convex. This note formulates a dual to the problem of minimizing a convex differentiable function under non-linear, concave constraints; this dual problem reduces to theirs in the case of linear constraints.

It has not been possible to establish (b) and (c) in full, and there is evidence that they do not hold as stated. Theorem 2 below derives the existence of a solution to the dual from that for the primal, and Theorem 3 establishes (c), in part, in the case of linear constraints. Thus the problem studied here does not seem to enjoy the complete symmetry of the notion of duality which is found in the linear case.

The known state of affairs regarding duality for this sort of problem is summarized in the table of Sec. 4. Several questions remain open. The utility of all this is discussed in Sec. 5.

2. The problems. Let f be a convex, differentiable function of $x = (x_1, \dots, x_n)$; for each $i = 1, \dots, m$, let $g_i(x)$ be a concave, differentiable function. For any function ϕ , let $\nabla\phi$ denote the gradient of ϕ ,

$$\left(\frac{\partial\phi}{\partial x_1}, \dots, \frac{\partial\phi}{\partial x_n} \right).$$

Primal problem Minimize $f(x)$ subject to

$$g_i(x) \geq 0, \quad i = 1, \dots, m. \quad (1)$$

Dual problem Maximize $f(x) - \sum_{i=1}^m u_i g_i(x)$ subject to (2)

$$\nabla f(x) = \sum_i u_i \nabla g_i(x), \quad u \geq 0. \quad (3)$$

*Received August 25, 1960; revised manuscript received February 20, 1961.

A *solution* of one of these problems will be a point, x or (x, u) respectively, which achieves the extreme value sought. The problems will not in general have solutions; about the only relation we can state in their absence is that of Theorem 1 below.

Theorem 1. Let V be the infimum of $f(x)$ under the constraints (1), and v be the supremum of (2) under the constraints (3). Then

$$v \leq V.$$

Proof. Let x^* satisfy $g_i(x^*) \geq 0$ (all i), and (x, u) satisfy the constraints (3). We then have the chain of relations

$$\begin{aligned} f(x^*) - f(x) &\geq \nabla f(x)(x^* - x) \\ &= \sum_i u_i \nabla g_i(x)(x^* - x) \\ &\geq \sum_i u_i [g_i(x^*) - g_i(x)] \\ &\geq -\sum_i u_i g_i(x); \end{aligned}$$

the first is due to the convexity of f ; the equality, to the constraints (3); the next inequality to the concavity of the g_i ; and the last inequality to the non-negativity of u_i and $g_i(x^*)$. It follows that

$$f(x^*) \geq f(x) - \sum_i u_i g_i(x),$$

which proves the theorem in the case that both sets of constraints are consistent. Allowing the convention $V = +\infty$ in case (1) are inconsistent, and $v = -\infty$ in case (3) are, the theorem follows.

3. The duality theorems. The "constraint qualification" of Kuhn and Tucker [6, p. 483] will be assumed from now on. Stated for the primal problem at hand, it runs

Let x satisfy (1), and let dx be any vector differential such that $\nabla g_i(x) dx \geq 0$ for all i such that $g_i(x) = 0$; then dx is tangent to some arc contained in the set of all x satisfying (1).

The qualification serves to rule out certain singularities which might otherwise occur on the boundary of the constraint set. The constraints g_i will meet the qualification if, for example, the constraint set has an interior point relative to the convex determined by those constraints which are linear; and thus a set of entirely linear constraints meets it.

Under the constraint qualification and the properties that have been assumed for f and g_i , the Equivalence Theorem of Kuhn and Tucker [6, Theorem 3] reads

The point x^0 is a solution of the primal problem if, and only if, x^0 and some u^0 constitute a saddle point for the Lagrangian function

$$L(x, u) = f(x) - \sum_i u_i g_i(x), \tag{4}$$

that is,

$$L(x^0, u) \leq L(x^0, u^0) \leq L(x, u^0) \tag{5}$$

for all $u \geq 0$ and all x .

Our principal duality theorem is

Theorem 2. If x^0 solves the primal problem, then there exists u^0 so that (x^0, u^0) solves the dual problem, and the extrema are equal.*

Proof. The function $L(x, u)$ is convex in x for any $u \geq 0$. It consequently satisfies the following relation, which asserts that its graph as a function of x lies above any of its tangent planes:

$$L(y, u) - L(x, u) \geq \nabla_x L(x, u)(y - x).$$

Now $\nabla_x L(x, u) = 0$ for any (x, u) satisfying (3), so that if (x^1, u) and (x^2, u) both satisfy (3) we have both $L(x^1, u) - L(x^2, u) \geq 0$ and the reverse inequality, whence $L(x^1, u) = L(x^2, u)$; in other words, $L(x, u)$ is independent of x for (x, u) satisfying (3). Consequently,

$$\begin{aligned} L(x^0, u^0) &= \text{Max } \{L(x^0, u) \mid u \geq 0\} \\ &\geq \text{Max } \{L(x^0, u) \mid (x^0, u) \text{ satisfy (3)}\} \\ &= \text{Max } \{L(x, u) \mid (x, u) \text{ satisfy (3)}\} \\ &\geq L(x^0, u^0), \end{aligned}$$

so that $L(x^0, u^0)$ is the maximal objective value for the dual problem. This proves the first part of the theorem.

Finally

$$L(x^0, u^0) = f(x^0) - \sum_i u_i^0 g_i(x^0) = f(x^0),$$

because each $u_i g_i(x^0) = 0$; if this last statement did not hold we would have $u_i > 0$ and $g_i(x^0) > 0$ for some i , and then $L(x^0, u^0)$ could be increased by decreasing u_i^0 , in contradiction to the saddle-point property (5). The proof is complete.

Theorem 3. If the primal problem has only linear constraints and these are inconsistent, and the constraints of the dual problem are consistent, then the supremum of the objective function (2) of the dual problem is $+\infty$.

Proof. Let (x, u) satisfy the constraints (3) of the dual problem, and let $g_i(x) = A^i x - b_i$ (all i). Since these latter constraints are inconsistent, there exists [4, Lemmas 4 and 5] $u^* \geq 0$ such that $u^* A^i = 0$ (all i) and $\sum_i u_i b_i > 0$. It follows that the point $(x, u + tu^*)$ satisfies the constraints (3) for all t , and that $f(x) - \sum_i (u_i + tu_i^*) g_i(x) \rightarrow +\infty$ as $t \rightarrow +\infty$.

4. The state of affairs. The following four possibilities exist for both the primal and the dual problem: (1) the problem has a solution; (2) the constraints of the problem are consistent and the function to be extremized is bounded in the direction of extremization, but the problem has no solution—the bound is not attained; (3) the constraints are consistent but the function is not bounded in the direction of extremization; (4) the constraints are inconsistent.

The table below summarizes the known results in terms of the possibility of a given pair of outcomes for the two problems. Theorems 1 and 2 account for the impossible outcomes; examples exist for the outcomes known to be possible. Three cases are un-

*I am indebted to W. S. Dorn and R. E. Gomory for pointing out an error in an earlier version of this theorem.

settled; an example of a problem whose dual was bounded but had no solution would help clear them up. With the exception of the second row and the second column, which would constitute impossible categories for linear programs, and the annoying example (6), the table is just the same as it would be if both problems were linear.

TABLE.
Outcomes of primal and dual problems.

DUAL	PRIMAL			
	Has solution	Bounded, no solution	Unbounded	Inconsistent
Has solution	YES (5)	?	NO (1)	YES (6, 3)
Bounded, no soln.	NO (2)	?	NO (1)	?
Unbounded	NO (1, 2)	NO (1)	NO (1)	YES (5)
Inconsistent	NO (2)	YES (4)	YES (5)	YES (5)

- (1) Established by Theorem 1.
- (2) Established by Theorem 2.
- (3) "NO" if the constraints are linear, by Theorem 3.
- (4) Illustrated by the example: $m = 1, n = 1, f(x) = e^x, g(x) = 0$.
- (5) Illustrated by readily constructed linear programming problems [4, pp. 57, 58].
- (6) Illustrated by the example: $m = 1, n = 1, f(x) = 0, g(x) = -e^x$.

5. Critique. Suppose that the constraints of the primal problem are linear

$$g_i(x) = A_i x - b_i \quad (i = 1, \dots, m),$$

where A_i is a row of the m by n matrix A , and b_i is the i th component of the vector b . The constraints may then be written

$$Ax \geq b.$$

The dual problem assumes the form

Maximize $f(x) - uAx + ub$ subject to

$$\nabla f(x) = uA, \quad u \geq 0.$$

(Dorn [2] gives the objective function in the equivalent form

$$f(x) - \nabla f(x)x + ub.) \tag{6}$$

In the case of quadratic programming with linear constraints,

$$f(x) = px + \frac{1}{2}xQx,$$

where Q is an n by n symmetric matrix. The objective function for the dual problem becomes

$$px + \frac{1}{2}xQx - uAx + ub = -\frac{1}{2}xQx + ub,$$

and its constraints

$$p + xQ = uA.$$

Finally, of course, in the case of linear programming—quadratic programming with $Q = 0$ —we obtain the classical dual problem

Maximize ub subject to

$$p = uA, \quad u \geq 0.$$

A basic difference between the non-linear duality results and those for linear programming is that in the non-linear case the function f of the primal problem appears not only in the constraints of the dual, as expected, but remains involved in its objective function as well. This fact seems to diminish somewhat the usefulness of the dual problem for computational purposes: the primal should commonly be easier to solve. (However, the “cutting plane” method [5] for non-linear problems, formulated in dual terms for computational purposes, turns out to be effectively an algorithm for the direct solution of the dual non-linear problem.)

In certain cases the explicit appearance of x in the objective of the dual problem can be transformed away. If the mapping $y = \nabla f(x)$ is invertible (which will be the case if f is strictly convex), so that $x = h(y)$ may be written, then the dual problem in the case of linear constraints assumes the form

Maximize $f(h(y)) - yh(y) + ub$ subject to

$$y = uA, \quad u \geq 0.$$

Dennis [1] states the dual problem in this way, using the “Legendre transform” $f(h(y)) - yh(y)$ of the function f . His formulation is more attractive when f is quadratic. The invertibility of f entails the non-singularity of Q . Letting R be the transpose of the inverse of Q , under the substitution of y for x the dual of the quadratic problem can be stated

Maximize $-\frac{1}{2}yRy + pRy + ub - \frac{1}{2}pRp$ subject to

$$y = uA, \quad u \geq 0.$$

It should be noticed that the “nice” properties regarding convexity that have been postulated for the primal are not inherited by the dual in the general case, or even in the case of linear constraints. The original form (2) for the objective of the dual problem is a convex function of x (one does not like to maximize such a function), while its form (6) is neither convex nor concave; the set of (x, u) satisfying the dual constraints is not convex; and while the objective of the dual depends only upon u for points of this set, it is neither a convex nor a concave function of u .

Things are much better, however, in the quadratic-linear case. The constraints are linear and the objective is concave. The linearity of the constraints permits linear computations in handling the dual, and makes possible an entirely linear algorithm for

the solution of this problem. In the "simplex method for quadratic programming" [7], a point (x, u) satisfying the constraints of the dual is found for which x satisfies the constraints of the primal and such that the difference, $u(Ax - b)$, between the objective functions of the two problems vanishes. By Theorem 1, such a point solves the problem.

REFERENCES

1. Jack B. Dennis, *Mathematical programming and electrical networks*, Technology Press, Cambridge, Mass., 1959
2. William S. Dorn, *Duality in quadratic programming*, Quart. Appl. Math. **18**, No. 2, 155-162 (July 1960)
3. William S. Dorn, *A duality theorem for convex programs*, IBM J. Research and Development **4**, No. 4, 407-413 (Oct. 1960)
4. A. J. Goldman and A. W. Tucker, *Theory of linear programming*, Linear Inequalities and Related Systems (Annals of Mathematics Study, No. 38), H. W. Kuhn and A. W. Tucker (eds.), Princeton University Press, 1956, pp. 53-97
5. James E. Kelley, Jr., *The cutting-plane method for solving convex programs*, J. Soc. Ind. and Appl. Math. **8**, No. 4, 703-712 (Dec. 1960)
6. H. W. Kuhn and A. W. Tucker, *Nonlinear programming*, Proc. 2nd Berkeley Symp. on Mathematical Statistics and Probability, University Of California Press, 1951, pp. 481-492
7. Philip Wolfe, *The simplex method for quadratic programming*, Econometrica **27**, No. 3, 382-398 (July 1959)