

—NOTES—

DUALITY IN NONLINEAR PROGRAMMING*

By O. L. MANGASARIAN (*Shell Development Company*)

Abstract. The main result of this work is a converse theorem to a duality theorem for nonlinear programming recently established by Wolfe. The conditions on the present theorem are slightly stronger than those imposed by Wolfe. Hanson has given a different proof of a similar theorem but without stating in the theorem some important assumptions made in the proof. †

Introduction. Duality principles relate two programming problems one of which, the *primal*, is a constrained minimization (maximization) problem and the other, the *dual*, is a constrained maximization (minimization) problem in such a way that the existence of a solution to one of these problems insures a solution to the other and the extrema of the two problems are equal.

Duality in linear programming has been known for some time [4] and extensive use has been made of it in theoretical and computational applications. Only recently, however, has duality in nonlinear programming been investigated [1], [2], [3], [7], [5]. There have been only two papers [7], [5]† that were concerned with the completely general nonlinear problem, i.e. the case where the nonlinearity occurred in both the objective function and the constraints. Wolfe [7] gives a theorem which insures that a solution to the dual problem exists if a solution to the primal problem exists and the two extrema are equal. However, he fails to give its converse. Hanson [5] gives a theorem similar to that of Wolfe's and its converse, however, in the proof of the converse the existence and the differentiability of the inverse of the gradient of the Lagrangian function is assumed. This rather important assumption does not appear in the statement of the theorem. In the present paper, the precise conditions under which this converse theorem holds are stated, and the proof given is considerably simpler than that of Hanson.

In what follows, matrix notation will be employed. Lower case Roman letters (with obvious exceptions) will denote column vectors, capital letters will represent matrices and Greek letters scalars. A prime will denote the transpose of a vector or matrix. The operators ∇ , ∇_u and ∇_v are the column vectors whose components are respectively

$$\left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right)', \quad \left(\frac{\partial}{\partial u_1}, \dots, \frac{\partial}{\partial u_m} \right)', \quad \left(\frac{\partial}{\partial v_1}, \dots, \frac{\partial}{\partial v_n} \right)'.$$

The constraints $g(x) \geq 0$ of the primal problem (1), (2) will be assumed to satisfy the Kuhn-Tucker constraint qualification ([6], p. 483).

At the point x^0 which is the solution of the primal problem, (1), (2) there will be associated with every component $g_i(x^0)$ of the primal constraint set a component u_i^0 of the Lagrange multiplier vector u^0 . The term *active primal constraints* will be used to

*Received December 7, 1961.

†After finishing this work, a forthcoming paper [8] was brought to the author's attention by W. S. Dorn. The converse duality theorem given therein is essentially that given here but the conditions under which it holds are more explicit here.

denote those components of $g_i(x^0)$ for which $u_i^0 \neq 0$. It can be shown [6] that for all such constraints $g_i(x^0) = 0$.

Wolfe's Duality Theorem ([7], Theorem 2): *If x^0 is a solution of the primal problem*

$$\text{Minimize } \varphi(x) \tag{1}$$

$$\text{subject to } g(x) \geq 0, \tag{2}$$

where φ is a differentiable, convex function of the n -vector x and each of the m components of $g(x)$ is a differentiable concave function of x , then there exists some m -vector $u^0 \geq 0$ such that (x^0, u^0) is the solution of the dual problem

$$\text{Maximize } \psi(x, u) \equiv \varphi(x) - u'g(x) \tag{3}$$

$$\text{subject to } \nabla\psi(x, u) \equiv \nabla\varphi(x) - \nabla u'g(x) = 0 \tag{4}$$

$$\text{and } u \geq 0. \tag{5}$$

Also,

$$\varphi(x^0) = \psi(x^0, u^0). \tag{6}$$

The Converse Duality Theorem. *Let (x^0, u^0) be a solution of the dual problem (3), (4), (5) where $\varphi(x)$ is a convex and twice continuously differentiable function of x and the components $g_i(x)$, $i = 1, \dots, m$ of $g(x)$ are concave and twice continuously differentiable functions of x . The vector x^0 is a solution of the primal problem (1), (2) and equation (6) holds if (in addition to the qualifications on $\varphi(x)$ and $g(x)$ already stated) either $\varphi(x)$ is strictly convex in the neighborhood of x^0 or if at least one of the active primal constraints $g_i(x^0)$ (i.e., those for which $u_i^0 \neq 0$) is strictly concave in the neighborhood of x^0 or both. If $\varphi(x)$ is quadratic and if $g(x)$ is linear then this converse theorem is true if $\varphi(x)$ is merely convex and twice differentiable.*

Proof: Define the function

$$\theta(x, u, v) \equiv \varphi(x) - u'g(x) + v'\nabla\varphi(x) - v'\nabla u'g(x), \tag{7}$$

where v is an n -by-1 column of variables. The necessary conditions for (x^0, u^0) to be solution of the dual problem (3), (4), (5) are given by Kuhn and Tucker ([6], Theorem 1 and Sect. 8, Case 2) as follows.* There must exist a v^0 such that

$$\nabla\theta(x^0, u^0, v^0) \equiv \nabla\varphi(x^0) - \nabla u^0'g(x^0) + \nabla v^0'\nabla\varphi(x^0) - \nabla v^0'\nabla u^0'g(x^0) = 0, \tag{8}$$

$$\nabla_u\theta(x^0, u^0, v^0) \equiv \nabla\varphi(x^0) - \nabla u^0'g(x^0) = 0, \tag{9}$$

$$\nabla_x\theta(x^0, u^0, v^0) \equiv -g(x^0) - \nabla_x v^0'\nabla u^0'g(x^0) \leq 0, \tag{10}$$

$$u^0'\nabla_x\theta(x^0, u^0, v^0) \equiv u^0'[-g(x^0) - \nabla_x v^0'\nabla u^0'g(x^0)] = 0, \tag{11}$$

$$u^0 \geq 0. \tag{12}$$

If we define the symmetric matrices R and Q^k , $k = 1, \dots, m$ whose respective ij th elements are $\partial^2\varphi(x^0)/\partial x_i\partial x_j$ and $\partial^2g_k(x^0)/\partial x_i\partial x_j$ then substitution from (9) in (8) gives

$$\left[R - \sum_{k=1}^m u_k^0 Q^k \right] v^0 = 0 \tag{13}$$

*To insure the validity of the necessary conditions (8) to (12) the dual constraints (4) must be such that the Jacobian of $\nabla\psi(x, u)$ with respect to x be different from zero at (x^0, u^0) or that the matrix (13a) below be non-singular. That this is indeed the case follows from the assumptions made in the theorem.

The assumptions made regarding the strict convexity of $\varphi(x)$ and/or one of the active primal constraints $g_i(x)$ together with the nonnegativity of u^0 assures that

$$R - \sum_{k=1}^m u_k^0 Q^k \tag{13a}$$

is a positive definite matrix and thus nonsingular. It follows from (13) that $v^0 = 0$. Conditions (9) through (12) become

$$\nabla\varphi(x^0) - \nabla u^0 g(x^0) = 0, \tag{14}$$

$$g(x^0) \geq 0, \tag{15}$$

$$u^0 g(x^0) = 0, \tag{16}$$

$$u^0 \geq 0. \tag{17}$$

But these conditions, together with the convexity of $\psi(x, u)$ are precisely the necessary and sufficient conditions for (x^0, u^0) to be a saddle point of $\psi(x, u)$ for $u \geq 0$ ([6] Theorems 1 and 2, and Sect. 8, Case 1), that is

$$\psi(x, u^0) \geq \psi(x^0, u^0) \geq \psi(x^0, u) \tag{18}$$

$$\text{for all } u \geq 0. \tag{19}$$

It follows then from the Equivalence Theorem ([6], Theorem 3, and Sect. 8, Case 1) that the conditions (18) and (19) are the necessary and sufficient conditions for x^0 to be a solution of the primal problem (1), (2). Equation (6) follows from (16).

The proof of the last part of the theorem for quadratic functions and linear constraints has been given in [2].

Different Constraints. It can be shown without great effort from the previous results that Wolfe's duality theorem and its converse apply also to the following set of dual problems:*

<i>Primal</i>	<i>Dual</i>
Min. $\varphi(x)$	Max. $\varphi(x) - u'g(x) - x'\nabla[\varphi(x) - u'g(x)]$
subject to $g(x) \geq 0$	subject to $\nabla[\varphi(x) - u'g(x)] \geq 0$
$x \geq 0$	$u \geq 0$

The same qualifications on the functions $\varphi(x)$ and $g(x)$ apply here as stated in the duality theorems. In addition to the dual constraints $\nabla[\varphi(x) - u'g(x)] \geq 0$ must satisfy the Kuhn-Tucker constraint qualification stated earlier.

REFERENCES

1. J. B. Dennis, *Mathematical programming and electrical networks*, Wiley, New York, N. Y., 1959
2. W. S. Dorn, *Duality in quadratic programming*, Q. Appl. Math., **18**, 155-162 (1960)
3. W. S. Dorn, *A duality theorem for convex programs*, IBM J. Res. Dev., **4**, 407-413 (1960)
4. D. Gale, H. W. Kuhn, A. W. Tucker, *Linear programming and the theory of games*, Chapter 19 of Activity Analysis of Production and Allocation, Wiley, New York, 1951.
5. M. A. Hanson, *A duality theorem in nonlinear programming with nonlinear constraints*, Austral. J. Statist., **3**, 64-72 (1961)
6. H. W. Kuhn and A. W. Tucker, *Nonlinear programming*, Proceedings, 2nd Berkeley Symp. in math. statist. and probab., Univ. California Press, 1951, pp. 481-492
7. P. Wolfe, *A duality theorem for nonlinear programming*, Q. Appl. Math., **19**, 239-244 (1961)
8. P. Huard, *Dual programs*, IBM J. Res. Dev., **6**, 137-139 (1962)

*Hanson [5] gives two other sets of dual problems which however do not seem to be valid in general without making some further assumptions in the duality theorems.