WIS-CS-72-159
Computer Sciences Department
University of Wisconsin
1210 West Dayton Street
Madison, Wisconsin 53706

Received August 24, 1972

SECOND AND HIGHER ORDER DUALITY IN NONLINEAR PROGRAMMING[1)]

by

O. L. Mangasarian[2)]

Computer Sciences Technical Report #159
October 1972

1) This research was supported by NSF Grant GJ362

2) Computer Sciences Department, University of Wisconsin, Madison, Wisconsin 53706

ABSTRACT

A dual problem associated with a primal nonlinear programming problem is presented that involves second derivatives of the functions constituting the primal problem. Duality results are derived for this pair of problems. More general dual problems are also presented, and duality results for these problems are also given.

1. INTRODUCTION

If we consider the nonlinear programming problem

1.1 $$\underset{x}{\text{minimize}} \; \{f(x) \mid g(x) \leqq 0\}$$

where f and g are differentiable functions from R^n into R and R^m respectively, and linearize both f and g around some arbitrary but fixed point $\bar{x}$ in R^n, we have the following linear program

1.2 $$\underset{p \in R^n}{\text{minimize}} \; \{f(\bar{x}) + \nabla f(\bar{x})p \mid g(\bar{x}) + \nabla g(\bar{x})p \leqq 0\}$$

where $\nabla f(\bar{x})$ denotes the $1 \times n$ gradient of f at $\bar{x}$ and $\nabla g(\bar{x})$ denotes the $m \times n$ Jacobian of g at $\bar{x}$. If we now take the dual of this linear program we obtain the dual linear program.

1.3 $$\underset{u \in R^m}{\text{maximize}} \; \{f(\bar{x}) + ug(\bar{x}) \mid -u\nabla g(\bar{x}) = \nabla f(\bar{x}),\ u \geqq 0\}$$

Finally if we let $\bar{x}$ be variable in this dual program we obtain

1.4 $$\underset{x \in R^n, u \in R^m}{\text{maximize}} \; \{f(x) + ug(x) \mid \nabla f(x) + u\nabla g(x) = 0,\ u \geqq 0\}$$

which is precisely the classical dual of nonlinear programming introduced, in a different way, by Wolfe [9] and investigated extensively [4] in the nonlinear programming literature.

Suppose now we repeat this process but with the following changes: Take quadratic instead of linear approximations of f and g around some fixed $\bar{x}$ (assuming that f and g are twice continuously differentiable) and take the dual of the resulting quadratic program. Thus we first take the following quadratic approximation to the original problem

$$1.5 \quad \underset{p\epsilon R^n}{\text{minimize}} \ \{f(\bar{x}) + \nabla f(\bar{x})p + \frac{1}{2}p\nabla^2 f(\bar{x})p \,|\, g_j(\bar{x}) + \nabla g_j(\bar{x})p + \frac{1}{2}p\nabla^2 g_j(\bar{x})p \leqq 0, \ j = 1,\ldots,m\}$$

where $\nabla^2 f(\bar{x})$ is the $n \times n$ symmetric Hessian at $\bar{x}$ and similarly for $\nabla^2 g_j(\bar{x})$. If we now take the dual of this program we obtain

$$1.6 \quad \underset{p\epsilon R^n, u\epsilon R^m}{\text{maximize}} \ \{f(\bar{x}) + \nabla f(\bar{x})p + \frac{1}{2}p\nabla^2 f(\bar{x})p + ug(\bar{x}) + u\nabla g(\bar{x})p + \frac{1}{2}p\nabla^2 ug(\bar{x})p \,|\, \nabla f(\bar{x}) + p\nabla^2 f(\bar{x}) + u\nabla g(\bar{x}) + p\nabla^2 ug(\bar{x}) = 0, \ u \geqq 0\}$$

where $\nabla^2 ug(\bar{x}) = \sum_{j=1}^{m} u_j \nabla^2 g_j(\bar{x})$. If we simplify the objective function by postmultiplying the first constraint by p and substituting in the objective function and if we let $\bar{x}$ be variable then we obtain

$$1.7 \quad \underset{x\epsilon R^n, p\epsilon R^n, u\epsilon R^m}{\text{maximize}} \ \{f(x) + ug(x) - \frac{1}{2}p(\nabla^2 f(x) + \nabla^2 ug(x))p \,|\, \nabla f(x) + u\nabla g(x) + p(\nabla^2 f(x) + \nabla^2 ug(x)) = 0, \ u \geqq 0\}$$

This is what we call the <u>second order dual</u> of the original nonlinear program 1.1.

If we define the Lagrangian function

$$L(x,u) = f(x) + ug(x)$$

then the second order dual can be written more simply as

1.7 $$\underset{x\in R^n, p\in R^n, u\in R^m}{\text{maximize}} \{L(x,u) - \frac{1}{2}p\nabla^2 L(x,u)p \,|\, \nabla L(x,u) + p\nabla^2 L(x,u) = 0$$

$$u \geqq 0\}$$

where we have used the notation which will be followed throughout the paper that $\nabla L(x,u)$ is the $1 \times n$ gradient with respect to the first argument x of L and similarly $\nabla^2 L(x,u)$ is the $n \times n$ Hessian matrix of L with respect to its first argument x. All gradient operators ∇ without a subscript will be gradients with respect to x or y. Gradients with respect to the second argument will be explicitly subscripted thus $\nabla_u L(x,u)$.

Many of the results of nonlinear programming duality [9,4] go through for the second order dual under appropriate conditions. The main condition, besides obvious ones on the second order terms, is an inclusion condition 2.8 or 2.9 that is needed only in the weak and forward duality theorems 2.1 and 2.14 but not in the converse duality theorem 2.20. The inclusion condition can be interpreted as a smallness requirement on either of the quantities $||p_2|| = (p_2 p_2)^{\frac{1}{2}}$ or $||x_1 - x_2||$

where x_1 is primal feasible for problem 1.1 and (x_2, p_2, u_2) is dual feasible for problem 1.7. At the optimum, it turns out that both of these quantities are zero.

We observe here that if we set $p = 0$, then 1.7 becomes the classical nonlinear programming dual [9], that is the first order dual 1.4. Hence it seems reasonable to say that p is a measure of the second order effects. In fact if we assume that $\nabla^2 L(x,u)$ is nonsingular problem 1.7 can be simplified as follows

1.8 $$\underset{x \in R^n, u \in R^m}{\text{maximize}} \; \{L(x,u) - \tfrac{1}{2}\nabla L(x,u)\nabla^2 L(x,u)^{-1}\nabla L(x,u) \mid u \geqq 0\}$$

where p has been eliminated through the constraint equality

1.9 $$p = -\nabla L(x,u)\nabla^2 L(x,u)^{-1}$$

Problem 1.8 is related to the subproblem of the dual feasible direction algorithm [5] for which fast numerical experience has been observed and for which quadratic convergence has been established for some special cases [3]. Problem 1.8 is also related to the dual of the subproblem of Wilson's algorithm [8] which Robinson has shown to converge quadratically [7]. In fact, it was consideration of the subproblems of these two algorithms that led to the formulation of the second order dual.

It is also possible to extend the concept of the second order dual by considering approximations other than quadratic around $\bar{x}$ and taking the first order dual of that problem. This is done in Section 3 of the paper which contains a duality theorem 3.3 and a converse duality theorem 3.6. The results for the second order duality are in Section 2 and consist of a weak duality theorem 2.1, a duality theorem 2.14 and a converse duality theorem 2.20. All the proofs are in Section 4.

2. SECOND ORDER DUALITY

We begin by establishing a weak duality relation between feasible points of the primal and of the second order dual. We emphasize here a distinguishing feature from the first order weak duality theorem [9,4] which is the inclusion condition 2.8 or 2.9. Without this inclusion condition the weak duality theorem 2.1 below does not hold but rather a part of the duality theorem 2.14 holds in which only a Kuhn-Tucker point of the second order dual 1.7 is guaranteed but not a maximum.

2.1 Weak Duality Theorem Let x be a primal feasible point and let (y,p,u) be a dual feasible point, that is

2.2 $g(x) \leqq 0$ (primal feasibility)

2.3 $\nabla L(y,u) + p\nabla^2 L(y,u) = 0$ (dual feasibility)

2.4 $u \geqq 0$

Let f and g be twice continuously differentiable on an open set containing the line segment $[x,y] = \{s \mid s = (1-\lambda)y + \lambda x,\ 0 \leqq \lambda \leqq 1\}$ and let

2.5 $z\nabla^2 L(s,u)\, z \geqq k(s)\, ||z||^2,\ \forall z \in R^n,\ \forall s \in [x,y]$ and some $k(s) > 0$

2.6 $||\nabla^2 L(y,u)|| \leqq K(y)$ for some $K(y) \geqq k(y) > 0$

Then the following inequality holds between the primal and dual objective functions

2.7 $f(x) \geqq L(y,u) - \frac{1}{2} p\nabla^2 L(y,u)p$

provided that the following inclusion condition holds

$$2.8 \qquad \|p\| \leqq \left(\frac{K(y)}{k(y)} - \left(\frac{K(y)^2}{k(y)^2} - \frac{k}{k(y)} \right)^{\frac{1}{2}} \right) \|x-y\|$$

or (inclusion condition)

$$2.9 \qquad \|p\| \geqq \left(\frac{K(y)}{k(y)} + \left(\frac{K(y)^2}{k(y)^2} - \frac{k}{k(y)} \right)^{\frac{1}{2}} \right) \|x-y\|$$

where

$$2.10 \qquad k = \inf_{s\in[x,y]} k(s) > 0 \quad \blacksquare$$

For the sake of easier readability we collect all proofs in Section 4.

We observe first that conditions 2.5 and 2.6 can be replaced by the slightly stronger but possibly simpler to verify conditions

$$2.11 \quad \begin{cases} z\nabla^2 f(s)z \geqq k_0(s) \|z\|^2, \ \forall z\in R^n, \ \forall s\in[x,y] \text{ and some } k_0(s) \\ z\nabla^2 g_j(s)z \geqq k_j(s) \|z\|^2, \ \forall z\in R^n, \ \forall s\in[x,y] \text{ and some } k_j(s), \ j=1,\ldots,m \\ k(s) = k_0(s) + \sum_{j=1}^{m} u_j k_j(s) > 0 \\ z\nabla^2 L(y,u)z \leqq K(y) \|z\|^2, \ \forall z\in R^n \text{ and some } K(y) \geqq k(y) > 0 \end{cases}$$

We also observe that none of the conditions 2.5 or 2.11 impose convexity requirements on <u>all</u> the functions f and g_j. Condition 2.5 is an n-dimensional uniform strict convexity requirement on $L(s,u)$ for all s

in the segment $[x,y]$. The first three conditions of 2.11 impose an n-dimensional uniform strict convexity requirement on some of the functions $f(s)$ and $g_j(s)$ for all s in the line segment $[x,y]$.

We remark also that the inclusion condition 2.8 holds automatically if $p = 0$, and similarly 2.9 holds automatically if $x-y = 0$. These two cases correspond respectively to reduction to the first order duality case,and to the case where the dual variably y is held fixed which is in effect problem 1.6. We formalize this as the following corollary.

2.12 Corollary The weak duality theorem 2.1 holds with the closeness conditions 2.8 and 2.9 replaced by

2.13 $$\|x-y\| \quad \|p\| = 0 \qquad \blacksquare$$

We proceed now to the duality theorem which relates a Kuhn-Tucker point or local maximum solution of the second order dual problem 1.7 with each local or global solution of the primal problem 1.1

2.14 Duality Theorem Let f and g be twice continuously differentiable on R^n and let $\bar{x}$ be a solution of 1.1 at which a constraint qualification is satisfied [4,p. 105]. Then $\bar{x}$, $\bar{p} = 0$ and some $\bar{u}$ satisfy the Kuhn-Tucker conditions for the second order dual problem 1.7 and the two objective functions are equal. In addition, $(\bar{x},\bar{p}=0,\bar{u})$ solve 1.7 under the following additional inclusion conditions

2.15 $$\|p\| \leqq \left(\frac{K(x)}{k(x)} - \left(\frac{K(x)^2}{k(x)^2} - \frac{k}{k(x)} \right)^{\frac{1}{2}} \right) \|x-\bar{x}\|$$

or

2.16 $$\|p\| \geqq \left(\frac{K(x)}{k(x)} + \left(\frac{K(x)^2}{k(x)^2} - \frac{k}{k(x)} \right)^{\frac{1}{2}} \right) \|x-\bar{x}\|$$

where $K(x)$, $k(x)$, and k satisfy, for all x in some set X in R^n containing $\bar{x}$ and for all u in some set U in R^m containing $\bar{u}$:

2.17 $z\nabla^2 L(s,u)z \geqq k(s) \|z\|^2 \quad \forall z \in R^n,\ \forall s \in X$ and some $k(s) > 0$.

2.18 $\|\nabla^2 L(x,u)\| \leqq K(x)$ for some $K(x) \geqq k(x) > 0$.

2.19 $k = \inf_{s \in X} k(s) > 0$ ▮

We observe here again that the inclusion condition 2.15 is automatically satisfied if $p = 0$, and 2.16 is automatically satisfied if $x = \bar{x}$.

The final result for the second order dual problem is a converse duality theorem which does not require an inclusion condition of the type 2.15 or 2.16.

2.20 <u>Converse Duality Theorem</u> Let f and g be three times differentiable on R^n and let $(\bar{x},\bar{p},\bar{u})$ be a local or global solution of the second order dual 1.7 or let $(\bar{x},\bar{p},\bar{u})$ satisfy the Fritz John or Kuhn-Tucker necessary optimality conditions for 1.7 [6,4 p. 170]. If $\nabla^2 L(\bar{x},\bar{u})$ is nonsingular and if any of the n, $n \times n$ matrices: $\frac{\partial}{\partial x_k} \nabla^2 L(\bar{x},\bar{u})$, $k=1,\ldots,n$, is positive or negative definite, then $(\bar{x},\bar{u})$ satisfy the Kuhn-Tucker conditions for the primal problems 1.1, that is

$$\nabla L(\bar{x},\bar{u}) = 0 \ , \quad \bar{u}g(\bar{x}) = 0 \ , \ g(\bar{x}) \leqq 0, \ \bar{u} \geqq 0$$

and the two objective functions are equal. If in addition f and g are convex at $\bar{x}$, or if f is pseudoconvex at $\bar{x}$ and g is quasiconvex at $\bar{x}$, then x is a global solution of 1.1. ▮

The proof of this theorem, given in Section 4, employs the Fritz-John necessary optimality conditions in the presence of equalities which were developed by Fromovitz and the author [6,4, p. 170]. Craven and Mond [1] have given an elegant proof of the first order converse duality theorem using the same Fritz-John conditions.

We note here that the assumption 2.17 of the duality theorem 2.14 implies that $\nabla^2 L(\bar{x},\bar{u})$ is positive definite and hence the second order sufficiency conditions of nonlinear programming are satisfied [2, p. 30]. Similarly the assumptions of the converse duality theorem 2.20 that $\nabla^2 L(\bar{x},\bar{u})$ is nonsingular and the convexity of f and g imply that $\nabla^2 L(\bar{x},\bar{u})$ is positive definite and hence again the second order sufficency conditions are satisfied.

3. HIGHER ORDER DUALITY

The results of the previous section can be extended by taking approximations of f and g around $\bar{x}$ more general than quadratic and then taking the nonlinear dual of the resulting problems. We thus obtain what we have termed here as a higher order dual. In particular consider problem 1.1 again and take nonlinear approximations of f and g around some arbitrary but fixed point $\bar{x}$. The approximations of f and g are given by $f(\bar{x}) + h(\bar{x},p)$ and $g(\bar{x}) + k(\bar{x},p)$ respectively where h maps $R^n \times R^n$ into R, k maps $R^n \times R^n$ into R^m and both of which are differentiable and satisfy certain assumptions to be stated below. (For the first order dual $h(x,p) = \nabla f(x)p$ and $k(x,p) = \nabla g(x)p$. For the second order dual $h(x,p) = \nabla f(x)p + \frac{1}{2}p\nabla^2 f(x)p$ and $k_j(x,p) = \nabla g_j(x)p + \frac{1}{2}p\nabla^2 g_j(x)p$, $j=1,\ldots,m$.) With the approximations $f(\bar{x}) + h(\bar{x},p)$ to $f(x)$ and $g(\bar{x}) + k(\bar{x},p)$ to $g(x)$ problem 1.1 is replaced by the following nonlinear program

3.1 $$\underset{p\in R^n}{\text{minimize}} \ \{f(\bar{x}) + h(\bar{x},p) \mid g(\bar{x}) + k(\bar{x},p) \leqq 0\}$$

Now taking the nonlinear dual of 3.1 we have

$$\underset{p\in R^n, u\in R^m}{\text{maximize}} \ \{f(\bar{x}) + h(\bar{x},p) + ug(\bar{x}) + uk(\bar{x},p) \mid \nabla_p h(\bar{x},p) + \nabla_p uk(\bar{x},p) = 0,\ u \geqq 0\}$$

where $\nabla_p h(\bar{x},p)$ denotes the $1 \times n$ gradient of $h(\bar{x},p)$ with respect to p and $\nabla_p uk(\bar{x},p)$ denotes the $1 \times n$ gradient of $uk(\bar{x},p)$ with respect to p . If we now remove the restriction on $\bar{x}$ and let it be a free variable we obtain the higher order dual problem

3.2 $$\underset{x\in R^n, p\in R^n, u\in R^n}{\text{maximize}} \quad \{f(x) + h(x,p) + ug(x) + uk(x,p) \mid \nabla_p h(x,p) + \nabla_p uk(x,p) = 0, \quad u \geqq 0\}$$

For this general higher order dual, it is not very simple to state conditions for the weak duality and duality theorems to hold. Instead we shall give a limited version of the duality theorem and a full version of the converse duality theorem. The limitation of the duality theorem consists in that a Kuhn-Tucker point of 3.2 is guaranteed rather than a maximum.

3.3 Duality Theorem Let f and g be differentiable on R^n, and let h and k be differentiable on $R^n \times R^n$. Let $\bar{x}$ be a local or global solution of 1.1 such that a constraint qualification is satisfied at $\bar{x}$ and let

3.4 $$\nabla h(\bar{x},0) = 0, \quad \nabla k(\bar{x},0) = 0$$

3.5 $$h(\bar{x},0) = 0,\ k(\bar{x},0) = 0,\ \nabla_p h(\bar{x},0) = \nabla f(\bar{x}),\ \nabla_p k(\bar{x},0) = \nabla g(\bar{x})$$

Then $\bar{x}$, $\bar{p} = 0$ and some $\bar{u}$ satisfy the Kuhn-Tucker conditions for the higher order dual problem 3.2 and the two objective functions are equal at these points. ∎

We observe here that the Kuhn-Tucker conditions for the higher order dual are given by conditions 4.15 to 4.21 with $\bar{v}_0 = 1$. We also note that conditions 3.4 and 3.5 are satisfied by both the linear and quadratic approximations from which the first and second order dual problems are obtained.

3.6 <u>Converse Duality Theorem</u> Let f and g be differentiable on R^n and let h and k be twice continuously differentiable on $R^n \times R^n$. Let $(\bar{x},\bar{p},\bar{u})$ be a local or global solution of the higher order dual problem 3.2 or let $(\bar{x},\bar{p},\bar{u})$ satisfy the Fritz John or Kuhn-Tucker conditions for 3.2. Let h and k satisfy 3.5 and the following conditions

3.7 $\nabla_p^2 (h(\bar{x},\bar{p}) + \bar{u}k(\bar{x},\bar{p}))$ is nonsingular

3.8 $$\left.\begin{array}{r} \nabla f(\bar{x}) + \nabla \bar{u}g(\bar{x}) + \nabla h(\bar{x},\bar{p}) + \nabla \bar{u}k(\bar{x},\bar{p}) = 0 \\ \nabla_p h(\bar{x},\bar{p}) + \nabla_p \bar{u}k(\bar{x},\bar{p}) = 0 \end{array}\right\rangle \Longrightarrow \bar{p} = 0$$

Then, $(\bar{x},\bar{u})$ satisfy the Kuhn-Tucker conditions for the primal problem 1.1 and the two objective functions are equal. If in addition f and g are convex at $\bar{x}$, or if f is pseudoconvex at $\bar{x}$ and g is quasiconvex at $\bar{x}$, then $\bar{x}$ is a global solution of 1.1 ∎

We remark here that the quadratic approximations satisfy the conditions 3.4,3.5,3.7 and 3.8 under the assumptions of the second order converse duality theorem 2.20.

4. PROOFS

Before proving the main results we need the following simple lemma.

4.1 Lemma Let h be a function mapping R^n into R, let x and y be points in R^n and let h be twice continuously differentiable on an open set containing the line segment $[x,y]$ and let

$$(y-x)\nabla^2 h(s)(y-x) \geqq k(s)\|y-x\|^2, \ \forall s\in[x,y] \text{ and some } k(s) > 0$$

Then

$$h(x) - h(y) - \nabla h(y)(x-y) \geqq \frac{k}{2}\|x-y\|^2$$

where

$$k = \inf_{s\in[x,y]} k(s)$$

Proof: Define the twice continuously differentiable function θ on $[0,1]$ by $\theta(\lambda) = h(y+\lambda(x-y))$. Then

$$\begin{aligned}
h(x) - h(y) &= \theta(1) - \theta(0) \\
&= \theta'(0) + \int_0^1 (1-\lambda)\theta''(\lambda)d\lambda \qquad \text{(by integration by parts)} \\
&= \nabla h(y)(x-y) + \int_0^1 (1-\lambda)(x-y)\nabla^2 h(y+\lambda(x-y))(x-y)d\lambda \\
&\geqq \nabla h(y)(x-y) + \int_0^1 (1-\lambda)k\|x-y\|^2 d\lambda \\
&= \nabla h(y)(x-y) + \frac{k}{2}\|x-y\|^2
\end{aligned}$$

Q.E.D.

Proof of Theorem 2.1

$$f(x) - (L(y,u) - \frac{1}{2}p\nabla^2 L(y,u)p)$$

$$= L(x,u) - ug(x) - L(y,u) + \frac{1}{2}p\nabla^2 L(y,u)p$$

$$\geqq \nabla L(y,u)(x-y) + \frac{k}{2}\|x-y\|^2 + \frac{k(y)}{2}\|p\|^2$$

(by 2.4, 2.2, 2.5, and 4.1)

$$= -p\nabla^2 L(y,u)(x-y) + \frac{k}{2}\|x-y\|^2 + \frac{k(y)}{2}\|p\|^2 \qquad \text{(by 2.3)}$$

$$\geqq -K(y)\|p\|\ \|x-y\| + \frac{k}{2}\|x-y\|^2 + \frac{k(y)}{2}\|p\|^2 \qquad \text{(by 2.6)}$$

$$= \frac{k(y)}{2}\left(\|p\| - \|x-y\|\left(\frac{K(y)}{k(y)} + \left(\frac{K(y)^2}{k(y)^2} - \frac{k}{k(y)}\right)^{\frac{1}{2}}\right)\right)$$

$$\left(\|p\| - \|x-y\|\left(\frac{K(y)}{k(y)} - \left(\frac{K(y)^2}{k(y)^2} - \frac{k}{k(y)}\right)^{\frac{1}{2}}\right)\right)$$

$$\geqq 0 \quad \text{(by 2.8 or 2.9)}$$

Q.E.D.

Proof of Theorem 2.14 Since a constraint qualification is satisfied at $\bar{x}$, it follows [4, p. 105] that $\bar{x}$ and some $\bar{u}$ in R^m satisfy the Kuhn-Tucker conditions

$$\nabla L(\bar{x},\bar{u}) = 0,\ \bar{u}g(\bar{x}) = 0,\ g(\bar{x}) \leqq 0 \quad \bar{u} \geqq 0 .$$

Hence $(\bar{x},\bar{p}=0,\bar{u})$ satisfy the Kuhn-Tucker condition for problem 1.7 (which are conditions 4.2 to 4.8 below with $\bar{v}_0 = 1$, $\bar{v} = 0$ and $\bar{w} = -g(\bar{x})$). The two objective functions are equal because $\bar{u}g(\bar{x}) = 0$ and $\bar{p} = 0$. We also have that $(\bar{x},\bar{p}=0,\bar{u})$ satisfy the constraints of 1.7 and the inclusion conditions 2.15 and 2.16. Let (x,p,u) satisfy the constraints of 1.7 and 2.15 or 2.16. Then

$$L(\bar{x},\bar{u}) - \frac{1}{2}\bar{p}\nabla^2 L(\bar{x},\bar{u})\bar{p}$$

$$= f(\bar{x}) \qquad (\text{since } \bar{u}g(\bar{x}) = 0 \text{ and } \bar{p} = 0)$$

$$\geqq L(x,u) - \frac{1}{2}p\nabla^2 L(x,u)p \qquad (\text{by theorem 2.1})$$

The inequality between the first and third terms above establish the theorem. Q.E.D.

Proof of Theorem 2.20 Since (x,p,u) constitute a local solution of 1.7, there exist $(\bar{v}_0,\bar{v},\bar{w})$ in $R \times R^n \times R^m$ not identically zero [6,4, p. 170] such that the following Fritz John conditions are satisfied

4.2 $\bar{v}_0\nabla(L(\bar{x},\bar{u})-\frac{1}{2}\bar{p}\nabla^2 L(\bar{x},\bar{u})\bar{p}) + \bar{v}\nabla^2 L(\bar{x},\bar{u}) + \nabla(\bar{p}\nabla^2 L(\bar{x},\bar{u})\bar{v}) = 0$

4.3 $-\bar{v}_0\nabla^2 L(\bar{x},\bar{u})\bar{p} + \nabla^2 L(\bar{x},\bar{u})\bar{v} = 0$

4.4 $\bar{v}_0(g_j(\bar{x})-\frac{1}{2}p\nabla^2 g_j(\bar{x})\bar{p}) + \nabla g_j(\bar{x})\bar{v} + \bar{p}\nabla \bar{g}_j^2(\bar{x})\bar{v} + \bar{w}_j = 0$

$j=1,\ldots,m$

4.5 $\bar{w}\bar{u} = 0$

4.6 $\nabla L(\bar{x},\bar{u}) + \bar{p}\nabla^2 L(\bar{x},\bar{u}) = 0$

4.7 $\bar{u} \geqq 0$

4.8 $(\bar{v}_0,\bar{w}) \geqq 0$

Since $\nabla^2 L(\bar{x},\bar{u})$ is nonsingular, 4.3 gives

4.9 $\bar{v} = \bar{v}_0\bar{p}$

If $\bar{v}_0 = 0$, then $\bar{v} = 0$ by 4.9 and $\bar{w} = 0$ by 4.4, but this contradicts the fact that $(\bar{v}_0,\bar{v},\bar{w}) \neq 0$. Hence

4.10 $\bar{v}_0 > 0$

Substitution of 4.9 in 4.2 gives

$$\bar{v}_0 \nabla(L(\bar{x},\bar{u}) - \frac{1}{2}\bar{p}\nabla^2 L(\bar{x},\bar{u})\bar{p}) + \bar{v}_0\bar{p}\nabla^2 L(\bar{x},\bar{u}) + \bar{v}_0\nabla(\bar{p}\nabla^2 L(\bar{x},\bar{u})\bar{p}) = 0$$

which in view of 4.10 and 4.6 gives

$$\frac{1}{2}\nabla(\bar{p}\nabla^2 L(\bar{x},\bar{u})\bar{p}) = 0$$

This is equivalent to

$$\sum_{i,j=1}^{i,j=n} \frac{\partial \nabla^2 L(\bar{x},\bar{u})_{ij}}{\partial x_k} \bar{p}_i \bar{p}_j = 0 \ , \ k=1,\ldots,n$$

where the ij subscripts denote component ij of the subscripted matrix $\nabla^2 L(\bar{x},\bar{u})$. But by assumption at least one of the n, $n \times n$ matrices, $\frac{\partial}{\partial x_k} \nabla^2 L(\bar{x},\bar{u})$, $k=1,\ldots,n$, is either positive or negative definite. Hence

4.11 $\quad \bar{p} = 0$ and $\bar{v} = \bar{v}_0\bar{p} = 0$.

Substitution of 4.11 and 4.4 in 4.5 and taking 4.10 into account gives

4.12 $\qquad \bar{u}g(\bar{x}) = 0$

Substitution of 4.11 in 4.4 and taking 4.8 and 4.10 into account gives

4.13 $\qquad g(\bar{x}) \leqq 0$

Setting $\bar{p} = 0$ in 4.6 gives

4.14 $\qquad \nabla L(\bar{x},u) = 0$

Conditions 4.14, 4.12, 4.13 and 4.7 are the Kuhn-Tucker condition for the primal problem 1.1 . The two objective functions are equal because $\bar{u}g(\bar{x}) = 0$ and $\bar{p} = 0$. If in addition f and g are convex at $\bar{x}$, or if f is pseudoconvex at $\bar{x}$ and g is quasiconvex at $\bar{x}$, then $\bar{x}$ is a global solution of 1.1 [4,p.153] . Q.E.D.

Proof of Theorem 3.3 Since $\bar{x}$ solves problem 1.1 then [4, p. 105] $\bar{x}$ and some $\bar{u}\in R^m$ satisfy the Kuhn-Tucker conditions

$$\nabla L(\bar{x},\bar{u}) = 0,\ \bar{u}g(\bar{x}) = 0,\ g(\bar{x}) \leqq 0,\ \bar{u} \geqq 0$$

These conditions together with 3.4 and 3.5 imply that $\bar{x}$, $\bar{p} = 0$, $\bar{u}$, $\bar{v} = 0$, $\bar{w} = -g(\bar{x})$, satisfy conditions 4.15 to 4.21 below with $\bar{v}_0 = 1$ which are the Kuhn-Tucker conditions for problem 3.2 . The two objective function are equal because $\bar{u}g(x) = 0$, $h(\bar{x},0) = 0$ and $k(\bar{x},0) = 0$. Q.E.D.

Proof of Theorem 3.6 Since $(\bar{x},\bar{p},\bar{u})$ is a solution of problem 3.2, then there exist nonzero $(\bar{v}_0,\bar{v},\bar{w})$ in $R \times R^n \times R^m$ [4, p. 170] such that the following Fritz John conditions hold

4.15 $\bar{v}_0\nabla(f(\bar{x})+h(\bar{x},\bar{p})+\bar{u}g(\bar{x})+\bar{u}k(\bar{x},\bar{p}))+\nabla(\nabla_p h(\bar{x},\bar{p})\bar{v}+\nabla_p \bar{u}k(\bar{x},\bar{p})\bar{v}) = 0$

4.16 $\bar{v}_0\nabla_p(h(\bar{x},\bar{p})+\bar{u}k(\bar{x},\bar{p}))+\bar{v}(\nabla_p^2 h(\bar{x},\bar{p})+\nabla_p^2 \bar{u}k(\bar{x},\bar{p})) = 0$

4.17 $\bar{v}_0 g(\bar{x})+\bar{v}_0 k(\bar{x},\bar{p})+\nabla_p k(\bar{x},\bar{p})\bar{v}+\bar{w} = 0$

4.18 $\bar{w}\bar{u} = 0$

4.19 $\nabla_p h(\bar{x},\bar{p})+\nabla_p \bar{u}k(\bar{x},\bar{p}) = 0$

4.20 $\bar{u} \geqq 0$

4.21 $(\bar{v}_0,\bar{w}) \geqq 0$

Now substitution of 4.19 into 4.16 and taking 3.7 into account gives

4.22 $\bar{v} = 0$

If $\bar{v}_0 = 0$, then by 4.17 and 4.22, $\bar{w} = 0$ and hence $\bar{v}_0 = 0$, $\bar{v} = 0$ and $\bar{w} = 0$

which is a contradiction. Hence $\bar{v}_0 \neq 0$ and by 4.21 we get that

4.23 $\bar{v}_0 > 0$

By 4.22, 4.23 and 4.15 we have

$$\nabla(f(\bar{x}) + h(\bar{x},\bar{p}) + \bar{u}g(\bar{x}) + \bar{u}k(\bar{x},\bar{p})) = 0$$

which together with 4.19 and 3.8 imply that

4.24 $\bar{p} = 0$

Substitution of 4.24 in 4.19 and noting 3.5 gives

4.25 $\nabla f(\bar{x}) + \nabla \bar{u}g(\bar{x}) = 0$

From 4.17, 4.21, 4.22, 4.23 and 3.5 we have that

4.26 $0 \geqq -\frac{\bar{w}}{\bar{v}_0} = g(\bar{x}) + k(\bar{x},0) = g(\bar{x})$

From 4.18, 4.23 and 4.26 we get that

4.27 $0 = -\frac{\bar{u}\bar{w}}{\bar{v}_0} = \bar{u}g(\bar{x})$

Conditions 4.25, 4.26, 4.27 and 4.20 are the Kuhn-Tucker conditions for problem 1.1. The two objective functions are equal because $\bar{u}g(\bar{x}) = 0$, $h(\bar{x},0) = 0$ and $k(\bar{x},0) = 0$. If in addition f and g are convex at $\bar{x}$ or if f is pseudoconvex at $\bar{x}$ and g is quasiconvex at $\bar{x}$, then $\bar{x}$ is a global solution of 1.1 [4, p. 153]. Q.E.D.

REFERENCES

1. B. D. Craven & B. Mond: "On Converse Duality in Nonlinear Programming," Operations Research, 19, 1971, 1075-1078.

2. A. V. Fiacco & G. P. McCormick: "Nonlinear Programming: Sequential Unconstrained Minimization Techniques," Wiley, New York, 1968.

3. U. Garcia-Palomares: "Superlinearly Convergent Algorithms for Nonlinear Programming," Ph.D. Dissertation, Computer Sciences Department, University of Wisconsin, Madison, 1972.

4. O. L. Mangasarian: "Nonlinear Programming," McGraw-Hill, New York, 1969.

5. O. L. Mangasarian: "Dual, Feasible Direction Algorithms," in "Techniques of Optimization," A. V. Balakrishnan (ed.), Academic Press, New York, 1972, 67-88.

6. O. L. Mangasarian & S. Fromovitz: "The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints," J. Math. Anal. Appl, 17, 1967, 37-47.

7. S. M. Robinson: "A Quadratically Convergent Algorithm for General Nonlinear Programming Problems," Report #1215, Math. Research Center, University of Wisconsin, Madison, 1972.

8. R. B. Wilson: "A Simplicial Method for Convex Programming," Ph.D. Dissertation, Harvard University, 1963.

9. P. Wolfe: "A Duality Theorem for Nonlinear Programming," Quart. Appl. Math. 19, 1961, 239-244.

BIBLIOGRAPHIC DATA SHEET	1. Report No. WIS-CS-72-159	2.	3. Recipient's Accession No.

4. Title and Subtitle Second and Higher Order Duality in Nonlinear Programming	5. Report Date October 1972
	6.
7. Author(s) Olvi L. Mangasarian	8. Performing Organization Rept. No. 159
9. Performing Organization Name and Address The University of Wisconsin Computer Sciences Dept., 1210 West Dayton St. Madison, Wisconsin 53706	10. Project/Task/Work Unit No.
	11. Contract/Grant No. NSF GJ362
12. Sponsoring Organization Name and Address National Science Foundation 1800 G. Street, N.W. Washington, D.C. 20550	13. Type of Report & Period Covered
	14.

15. Supplementary Notes

16. Abstracts

A dual problem associated with a primal nonlinear programming problem is presented that involves second derivatives of the functions constituting the primal problem. Duality results are derived for this pair of problems. More general dual problems are also presented, and duality results for these problems are also given.

17. Key Words and Document Analysis. 17a. Descriptors

Optimization, Duality, Mathematical Programming

17b. Identifiers/Open-Ended Terms

17c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
Unlimited	UNCLASSIFIED	22
	20. Security Class (This Page) UNCLASSIFIED	22. Price

FORM NTIS-35 (10-70) USCOMM-DC 40329-P71