COMPUTER SCIENCES
DEPARTMENT

University of Wisconsin-
Madison

LOCAL DUALITY OF NONLINEAR PROGRAMS

by

O. Fujiwara, S.-P. Han & O. L. Mangasarian

Computer Sciences Technical Report #463

January 1982

UNIVERSITY OF WISCONSIN-MADISON

LOCAL DUALITY OF NONLINEAR PROGRAMS

O. Fujiwara, S.-P. Han & O. L. Mangasarian

January 1982

ABSTRACT

It is shown that the second order sufficient (necessary) optimality condition for the dual of a nonlinear program is equivalent to the inverse of the Hessian of the Lagrangian being positive definite (semidefinite) on the normal cone to the local primal constraint surface. This compares with the Hessian itself being positive definite (semidefinite) on the tangent cone on the local primal constraint surface for the corresponding second order condition for the primal problem. We also show that primal second order sufficiency (necessity) and dual second order necessity (sufficiency) is essentially equivalent to the Hessian of the Lagrangian being positive definite. This follows from the following interesting linear algebra result: A necessary and sufficient condition for a nonsingular $n \times n$ matrix to be positive definite is that for each or some subspace of R^n, the matrix must be positive definite on the subspace and its inverse be positive semidefinite on the orthogonal complement of the subspace.

AMS (MOS) Subject Classifications: 90C30, 15A03

Key Words: Nonlinear programming, second order optimality, duality

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based on work supported by the National Science Foundation under Grants ENG-7903881 and MCS-7901066.

LOCAL DUALITY OF NONLINEAR PROGRAMS

O. Fujiwara, S.-P. Han & O. L. Mangasarian

1. Introduction

We consider the following nonlinear program

$$\text{(P)} \qquad \begin{array}{ll} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & g(x) \leqq 0 \\ & h(x) = 0 \end{array}$$

and its Wolfe dual [14,7]

$$\text{(D)} \qquad \begin{array}{ll} \underset{(x,u,v)}{\text{maximize}} & L(x,u,v) \\ \text{subject to} & \nabla_x L(x,u,v) = 0 \\ & u \geqq 0 \end{array}$$

where $f:R^n \to R$, $g:R^n \to R^m$ and $h:R^n \to R^q$ are differentiable functions on R^n, $L(x,u,v) := f(x) + u^T g(x) + v^T h(x)$ is the standard Lagrangian and $\nabla_x L$ is the gradient with respect to x. The relationships between the above two problems have been extensively studied for the convex case [14,7]. Our principal concern here are local duality results which in the absence of convexity assumptions require the use of second order optimality conditions.

In Section 2 we give a geometrically meaningful second order sufficient optimality condition in Definition 2.1 for the dual problem (D)

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based on work supported by the National Science Foundation under Grants ENG-7903881 and MCS-7901066.

and prove in Theorem 2.2 that it is equivalent to the standard second order sufficient optimality condition [6,2] applied to the dual problem (D). Thus, as is well known, the second order sufficient optimality condition for the primal problem (P) is the positive definiteness of the Hessian of the Lagrangian on the tangent cone to the local constraint surface. Our second order sufficient (necessary) optimality condition Theorem 2.2 (Theorem 2.5) for the dual problem (D) is that the inverse of the Hessian of the Lagrangian is positive definite (semidefinite) on the normal cone to the primal local constraint surface. It is worthwhile to note that while positive definiteness of the Hessian of the Lagrangian ensures the satisfaction of the second order sufficiency condition for the primal problem, this is not the case for the dual problem where a constraint qualification is needed (Theorem 2.3) in order to ensure that dual second order sufficiency holds under positive definiteness of the Hessian of the Lagrangian.

In Section 3 we characterize Karush-Kuhn-Tucker points of the primal problem that locally solve both the primal and dual problems simultaneously. We show (Theorems 3.6 and 3.7) that these points are essentially points where the Hessian of the Lagrangian with respect to the primal variables is positive definite. In order to establish these results we prove an interesting result of linear algebra (Theorem 3.5) which states that a necessary and sufficient condition for a nonsingular $n \times n$ matrix to be positive definite is that for each or some subspace S of R^n, A must be positive definite on S and A^{-1} must be positive semidefinite on the orthogonal complement $S^{\perp}$ of S.

We briefly describe our notation now. All vectors will be column vectors unless transposed to a row vector by the superscript T. For x in the n-dimensional real Euclidean space R^n, x_i, $i=1,\dots,n$, will denote its components. For an $m\times n$ real matrix we shall say that $A\in R^{m\times n}$, A_i will denote the ith row of A, and if $I\subset\{1,\dots,m\}$ then A_I will denote the submatrix with rows A_i, $i\in I$. For a differentiable function $g:R^n\to R^m$, $\nabla g(x)$ will denote the transpose of the $m\times n$ Jacobian matrix of g at x. For a twice differentiable function $L:R^{n+m}\to R$, $\nabla_x L(x,u)$ will denote the $n\times 1$ gradient with respect to x, $\nabla_u L(x,u)$ will denote the $m\times 1$ gradient with respect to u, $\nabla^2 L(x,u)$ will denote the $(n+m)\times(n+m)$ Hessian with respect to both x and u whose submatrix components are denoted as follows

$$\nabla^2 L(x,u) = \begin{bmatrix} \nabla_{xx}L(x,u) & \nabla_{xu}L(x,u) \\ \nabla_{ux}L(x,u) & \nabla_{uu}L(x,u) \end{bmatrix}.$$

2. Geometrically meaningful second order optimality condition for the dual problem

In order to establish local duality results without any convexity assumptions, second order necessary and sufficient optimality conditions become essential. The second order sufficient condition for the primal program (P) was given by McCormick and Fiacco [6,2] and has been extensively studied. Research on this topic continues (see, for example, [4,13]). In this section we formulate geometrically meaningful second order necessary and sufficient optimality conditions for the dual problem (D) and study the relationship to the corresponding conditions for the primal.

Recall that an $(n+m+q)$-vector $(\bar{x},\bar{u},\bar{v})$ is said to be a Karush-Kuhn-Tucker triple of the primal program (P) if the following conditions hold:

$$(2.1)\qquad \begin{array}{ll} \text{(a)} & \nabla_x L(\bar{x},\bar{u},\bar{v}) = 0 \\ \text{(b)} & g(\bar{x}) \leqq 0 \\ \text{(c)} & h(\bar{x}) = 0 \\ \text{(d)} & \bar{u} \geqq 0 \\ \text{(e)} & \bar{u}_i g_i(\bar{x}) = 0 \qquad i=1,\dots,m\,. \end{array}$$

Such a triple is said to satisfy the primal second order sufficient optimality condition if f, g and h are twice continuously differentiable at $\bar{x}$ and

$$(2.2)\qquad \left.\begin{array}{r} \nabla g_J(\bar{x})^T d = 0 \\ \nabla g_K(\bar{x})^T d \leqq 0 \\ \nabla h(\bar{x})^T d = 0 \\ d \neq 0 \end{array}\right\} \Longrightarrow d^T \nabla_{xx} L(\bar{x},\bar{u},\bar{v}) d > 0$$

where $J := \{i \mid g_i(\bar{x}) = 0 \text{ and } \bar{u}_i > 0\}$

$K := \{i \mid g_i(\bar{x}) = 0 \text{ and } \bar{u}_i = 0\}$.

We now give a second order sufficient optimality condition for the dual program (D) which we shall justify by Theorem 2.2 below.

Definition 2.1 A Karush-Kuhn-Tucker triple $(\bar{x},\bar{u},\bar{v})$ of the primal problem (P) is said to satisfy the dual second order sufficient optimality condition if f, g and h are twice continuously differentiable at $\bar{x}$, if the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ is nonsingular and

$$\text{(2.3)} \qquad \left.\begin{array}{ll} w = \nabla g(\bar{x})y + \nabla h(\bar{x})z & \\ y_i = 0 & i \in I \\ y_i \geqq 0 & i \in K \\ (y,z) \neq 0 & \end{array}\right\} \Longrightarrow w^T \nabla_{xx} L(\bar{x},\bar{u},\bar{v})^{-1} w > 0$$

where $I := \{i \mid g_i(\bar{x}) < 0\}$ and $K := \{i \mid g_i(\bar{x}) = 0,\ \bar{u}_i = 0\}$.

The geometric relationship between the primal and the dual second order sufficient conditions is an interesting one. Let T be a tangent cone of the local primal constraint surface at the point $\bar{x}$ induced by the second order optimality condition (2.2), that is,

$$\text{(2.4)} \qquad T := \{d \mid \nabla g_J(\bar{x})^T d = 0,\ \nabla g_K(\bar{x})^T d \leqq 0,\ \nabla h(\bar{x})^T d = 0\} .$$

Then the polar cone of T, denoted by N and called the normal cone at $\bar{x}$, is given by

$$\text{(2.5)} \qquad \begin{aligned} N &:= \{w \mid w^T d \leqq 0,\ \forall d \in T\} \\ &= \{w \mid w = \nabla g(\bar{x})y + \nabla h(\bar{x})z,\ y_I = 0,\ y_K \geqq 0\} . \end{aligned}$$

Therefore, the primal second order sufficient condition merely says that the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ is positive definite on the tangent cone T, while the dual second order sufficient condition says that the inverse Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})^{-1}$ is positive definite on the normal cone N. It will be shown in Section 3 that for both conditions to hold it is not only sufficient but also necessary that the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ be positive definite on the whole space R^n. We also note here that it was proved in [4] that the tangent cone can also be expressed as

$$T = \{d \mid \nabla f(\bar{x})^T d = 0,\ \nabla g_A(\bar{x})^T d \leqq 0,\ \nabla h(\bar{x})^T d = 0\}$$

where $A := \{i \mid g_i(\bar{x}) = 0\} = J \cup K$, that is A is the index set of all active inequality constraints at $\bar{x}$. Consequently, the normal cone N can also be written as

$$N = \{w \mid w = \mu\nabla f(\bar{x}) + \nabla g(\bar{x})y + \nabla h(\bar{x})z,\ y_A \geqq 0,\ y_I = 0\}\ .$$

These expressions contain the gradient $\nabla f(\bar{x})$ of the objective function and treat the index sets J and K on an equal footing.

We now justify Definition 2.1 by showing that the dual second order sufficient condition given in this definition is equivalent to the one derived by applying McCormick's second order sufficient optimality condition directly to the dual program (D).

Theorem 2.2 (Equivalence of dual second order sufficient optimality condition to McCormick's condition) If $(\bar{x},\bar{u},\bar{v})$ is a Karush-Kuhn-Tucker triple of the primal program (P) and if f, g and h are twice continuously differentiable at $\bar{x}$ then $(\bar{x},\bar{u},\bar{v})$ is a Karush-Kuhn-Tucker point of the dual program (D) with the (n+m)-vector $(0,-g(\bar{x}))$ as its Lagrange multiplier. Furthermore, the vector $(\bar{x},\bar{u},\bar{v},0,-g(\bar{x}))$ satisfies McCormick's second order sufficient optimality condition for (D) if and only if $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ is nonsingular and condition (2.3) holds.

Proof The first statement of the theorem follows immediately by direct verification. Notice that the Lagrangian for the dual program (D) is given by

$$M(x,u,v,s,t) := L(x,u,v) + s^T\nabla_x L(x,u,v) + t^T u$$

Let ∇M and $\nabla^2 M$ denote respectively the gradient and the Hessian of M with respect to (x,u,v) only. Thus we have

$$\nabla M(x,u,v,s,t) = \begin{bmatrix} \nabla_x L(x,u,v) + \nabla_{xx}L(x,u,v)s \\ g(x) + \nabla g(x)^T s + t \\ h(x) + \nabla h(x)^T s \end{bmatrix}.$$

Let $\bar{s} = 0$ and $\bar{t} = -g(\bar{x})$, then it follows that

$$\nabla^2 M(\bar{x},\bar{u},\bar{v},\bar{s},\bar{t}) = \begin{bmatrix} \nabla_{xx}L(\bar{x},\bar{u},\bar{v}) & \nabla g(\bar{x}) & \nabla h(\bar{x}) \\ \nabla g(\bar{x})^T & 0 & 0 \\ \nabla h(\bar{x})^T & 0 & 0 \end{bmatrix} \tag{2.6}$$

Therefore, McCormick's second order sufficient optimality condition [6,2] for problem (D) is that

$$\left.\begin{array}{l} \nabla_{xx}L(\bar{x},\bar{u},\bar{v})x + \nabla g(\bar{x})y + \nabla h(\bar{x})z = 0 \\ y_I = 0 \\ y_K \geqq 0 \\ (x,y,z) \neq 0 \end{array}\right\} \Longrightarrow (x^T,y^T,z^T)\nabla^2 M(\bar{x},\bar{u},\bar{v},\bar{s},\bar{t})\begin{bmatrix} x \\ y \\ z \end{bmatrix} < 0, \tag{2.7}$$

where as before the Hessian is with respect to (x,y,v) only,

$$I := \{i \mid g_i(\bar{x}) < 0\} \quad \text{and} \quad K := \{i \mid g_i(\bar{x}) = 0,\ \bar{u}_i = 0\}.$$

By (2.6) and the equality on the left-hand-side of (2.7), the inequality on the right-hand-side of (2.7) is equivalent to $x^T\nabla^2 L(\bar{x},\bar{u},\bar{v})x > 0$. Hence condition (2.7) can be expressed as follows

$$(2.8) \quad \left.\begin{array}{l} \nabla_{xx}L(\bar{x},\bar{u},\bar{v})x + \nabla g(\bar{x})y + \nabla h(\bar{x})z = 0 \\ y_I = 0 \\ y_K \geqq 0 \\ (x,y,z) \neq 0 \end{array}\right\} \Longrightarrow x^T\nabla_{xx}L(\bar{x},\bar{u},\bar{v})x > 0 .$$

We claim now that condition (2.8) implies the nonsingularity of the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$. Suppose it is not true, then there exists a nonzero $\hat{x} \in R^n$ such that $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})\hat{x} = 0$. Let $(\hat{y},\hat{z}) := (0,0)$, then $(\hat{x},\hat{y},\hat{z})$ satisfies the conditions in the left-hand-side of implication (2.8), and hence by implication (2.8) we have $\hat{x}^T\nabla_{xx}L(\bar{x},\bar{u},\bar{v})\hat{x} > 0$. This however contradicts $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})\hat{x} = 0$.

It now follows from the nonsingularity of $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ that the condition $(x,u,z) \neq 0$ in (2.8) can be replaced by $(y,z) \neq 0$, because $(y,z) = 0$ implies $x = 0$. Therefore, by defining $w := -\nabla_{xx}L(\bar{x},\bar{u},\bar{v})x$, condition (2.8) again can be rewritten as

$$(2.9) \quad \left.\begin{array}{l} w = \nabla g(\bar{x})y + \nabla h(\bar{x})z \\ y_I = 0 \\ y_K \geqq 0 \\ (y,z) \neq 0 \end{array}\right\} \Longrightarrow w^T\nabla_{xx}L(\bar{x},\bar{y},\bar{v})^{-1}w > 0 ,$$

which is condition (2.3) of our Definition 2.1.

Conversely, it is obvious that condition (2.9) and the nonsingularity of $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ imply (2.8) which is equivalent to McCormick's condition (2.7). □

It should be remarked here that because we used $\bar{s} = 0$ as a multiplier for the dual problem (D) we were able to get away without assuming that f, g and h are thrice differentiable at $\bar{x}$ but merely twice differentiable.

It is important to note that unlike the situation for the primal problem, where the second order sufficiency implication (2.2) holds automatically when $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ is positive definite, the second order sufficiency implication (2.3) for the dual problem need not hold when $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ is positive definite because $w = 0$ may satisfy the conditions of the left-hand-side of implication (2.3). However under a slightly more stringent version of the standard constraint qualification of nonlinear programming [9,7] we can show that positive definiteness of $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ does indeed ensure the satisfaction of the dual second order sufficiency implication (2.3) as follows.

<u>Theorem 2.3</u> (Dual second order sufficiency under positive definiteness of the Hessian of the Lagrangian) Let $(\bar{x},\bar{u},\bar{v})$ be a Karush-Kuhn-Tucker point of the primal problem (P), let f, g and h be twice continuously differentiable at $\bar{x}$, let $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ be positive definite and let the following primal constraint qualification hold at $\bar{x}$

$$(2.10)\quad \begin{cases} \nabla h_i(\bar{x}),\ i=1,\ldots,q,\ \nabla g_{i\in J}(\bar{x}), \text{ are linearly independent} \\ \text{and there exists a } p \in R^n \text{ such that} \\ \nabla g_K(\bar{x})^T p < 0,\ \nabla g_J(\bar{x})^T p = 0,\ \nabla h(\bar{x})^T p = 0 \end{cases}$$

Then the second order sufficiency implication (2.3) holds.

Proof Let (w,y,z) satisfy the conditions of the left-hand-side of implication (2.3). If $w \neq 0$ then implication (2.3) holds because $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})^{-1}$ is positive definite. We now show that if $w = 0$ we contradict the constraint qualification (2.10). If $y_K = 0$ then $(y_J, z) \neq 0$ and we contradict the linear independence of $\nabla h_i(\bar{x})$, $i=1,\ldots,q$, $\nabla g_{i\in J}(\bar{x})$. If $y_K \neq 0$ then we have the contradiction

$$0 = y_K^T \nabla g_K(\bar{x})^T p + y_J^T \nabla g_J(\bar{x})^T p + z^T \nabla h(\bar{x})^T p < 0. \qquad \square$$

Remark 2.4 It can be shown that the constraint qualification (2.10) implies the standard constraint qualification of nonlinear programming [9 (Definition 11.3.5), 7] and (2.10) itself is implied by the often used [12,8] linear independence assumption of all the active constraint gradients: $\nabla h_i(\bar{x})$, $i=1,\ldots,q$, $\nabla g_{i\in A}(\bar{x})$.

We now derive a second order necessary optimality condition for the dual problem which besides having a geometrically meaningful interpretation will be useful in characterizing simultaneous local solutions of the primal and dual problems. Recall that McCormick's second order necessary condition for the primal problem (P) is that the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ be positive semi-definite on the cone $\{d \mid \nabla g_A(\bar{x})^T d = 0,\ \nabla h(\bar{x})^T d = 0\}$ where $A := \{i \mid g_i(\bar{x}) = 0\}$. As expected, the second order necessary condition for the dual problem (D) is that the inverse Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})^{-1}$ be positive semi-definite on the normal cone N defined in (2.5). We give this result below. Note that, under our assumption of nonsingularity of the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$, no constraint qualification is required as is the case in McCormick's second order necessary condition.

Theorem 2.5 (Dual second order necessity) If $(\bar{x},\bar{u},\bar{v})$ is a local maximum point of the dual program (D), if f, g and h are twice continuously differentiable at $\bar{x}$, and if the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ is nonsingular, then

$$(2.11) \qquad \left.\begin{array}{ll} w = \nabla g(\bar{x})y + \nabla h(\bar{x})z & \\ y_i = 0 & i \in I \\ y_i \geqq 0 & i \in K \end{array}\right\} \Longrightarrow w^T \nabla_{xx} L(x,u,v)^{-1} w \geqq 0$$

where $I := \{i \mid g_i(\bar{x}) < 0\}$ and $K := \{i \mid g_i(\bar{x}) = 0 \text{ and } \bar{u}_i = 0\}$.

Proof Let vectors w, y and z be fixed vectors satisfying the conditions of the left-handed side of (2.11). We consider the function $F: R^{n+m+q+1} \to R^{n+m+q}$ defined by

$$F(x,u,v,t) := \begin{bmatrix} \nabla_x L(x,u,v) \\ u - \bar{u} - ty \\ v - \bar{v} - tz \end{bmatrix}$$

Clearly, we have $F(\bar{x},\bar{u},\bar{v},0) = 0$. Furthermore, it follows from the nonsingularity of $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ that the Jacobian $\nabla_{x,u,v}F(\bar{x},\bar{u},\bar{v},0)$ is also nonsingular. Hence, by the implicit function theorem, there exist a positive number ε and continuously differentiable functions $x(t)$, $u(t)$ and $v(t)$ defined on $(-\varepsilon,\varepsilon)$ such that $x(0) = \bar{x}$, $u(0) = \bar{u}$, $v(0) = \bar{v}$ and

(2.12)
$$\text{(a)}\quad \nabla_x L(x(t), u(t), v(t)) = 0$$
$$\text{(b)}\quad u(t) = \bar{u} + ty$$
$$\text{(c)}\quad v(t) = \bar{v} + tz \; .$$

Differentiating (2.12a) at $t = 0$, we get that

$$\nabla_{xx} L(\bar{x},\bar{u},\bar{v})x'(0) + \nabla g(\bar{x})y + \nabla h(\bar{x})z = 0 \, ,$$

which imples, for $w = \nabla g(\bar{x})y + \nabla h(\bar{x})z$, that

$$w = -\nabla_{xx} L(\bar{x},\bar{u},\bar{v})x'(0) \; .$$

Let $\theta(t) := L(x(t), u(t), v(t))$. Notice that for sufficiently small $t \in [0, \varepsilon)$ the vector $(x(t), u(t), v(t))$ is feasible to the dual program (D) and hence $\theta(0) \geqq \theta(t)$ for all sufficiently small nonnegative t. On the other hand, we have that

$$\theta'(0) = \nabla_x L(\bar{x},\bar{u},\bar{v})^T x'(0) + g(\bar{x})^T y + h(\bar{x})^T z = 0 \, .$$

Therefore it follows that $\theta''(0) \leqq 0$. By direct verification, we have that

$$\theta''(0) = -w^T \nabla_{xx} L(\bar{x},\bar{u},\bar{v})^{-1} w \, .$$

Hence, the proof is complete. □

3. Characterization of simultaneous local solutions of the primal and dual problems

In this section we characterize Karush-Kuhn-Tucker points of the primal problem that locally solve both the primal and dual programs simultaneously. We will show that these points are essentially those Karush-Kuhn-Tucker points at which the Hessian of the Lagrangian with respect to the primal variables is positive definite. To establish this we need a preliminary fundamental result, Theorem 3.5 below, which appears to be an interesting linear algebra result in its own right. Related results have appeared in [10]. First we recall a well known result.

Lemma 3.1 (Sherman-Morrison-Woodbury formula [11]) Let $A \in R^{n\times n}$ be nonsingular and let $U, V \in R^{n\times m}$. Then $A + UV^T$ is nonsingular if and only if $I + V^TA^{-1}U$ is nonsingular, in which case

$$(A+UV^T)^{-1} = A^{-1} - A^{-1}U(I+V^TA^{-1}U)^{-1}V^TA^{-1} .$$

We can now state and prove our preliminary fundamental result.

Theorem 3.2 (Characterization of positive definitness of a nonsingular matrix) A necessary and sufficient condition for a nonsingular matrix $A \in R^{n\times n}$ to be positive definite is that for some or each $\alpha > 0$ and $B \in R^{m\times n}$, the matrices $A + \alpha B^TB$ and $I + \alpha BA^{-1}B^T$ be positive definite.

Proof (Necessity) Let A be positive definite, then obviously for each $\alpha > 0$ and each $B \in R^{m\times n}$, the matrices $A + \alpha B^TB$ and $I + \alpha BA^{-1}B^T$ are positive definite.

(Sufficiency) Let A be nonsingular and let, for some $\alpha > 0$ and some $B \in R^{m+n}$, the matrices $A + \alpha B^T B$ and $I + \alpha B A^{-1} B^T$ be positive definite. By the Sherman-Morrison-Woodbury formula we have that

$$A^{-1} = (A+\alpha B^T B)^{-1} + \alpha A^{-1} B^T (I+\alpha B A^{-1} B^T)^{-1} B A^{-1}$$

which is the sum of a positive definite matrix and a positive semidefinite matrix and hence A^{-1} and consequently A is positive definite. □

By noting that if $I + \alpha B A^{-1} B^T$ is positive definite for all $\alpha \geqq \bar{\alpha}$ for some $\bar{\alpha}$ then $B A^{-1} B^T$ must be positive semidefinite the following corollary follows from Theorem 3.2.

<u>Corollary 3.3</u> (Alternate characterization of positive definiteness of a nonsingular matrix) Let $A \in R^{n\times n}$ be nonsingular. A is positive definite if and only for all $\alpha \geqq \bar{\alpha}$ for some $\bar{\alpha} > 0$ and for some or all $B \in R^{m\times n}$, $A + \alpha B^T B$ is positive definite and $B A^{-1} B^T$ is positive semidefinite.

A very nice geometric version of Corollary 3.3 can be stated if we make use of the following well known result of positive definitness of quadratic forms on a linear subspace.

<u>Lemma 3.4</u> (Finsler-Debreu Lemma [3,1]) Let $A \in R^{n\times n}$ and let $B \in R^{m\times n}$. Then

$$\left\langle \begin{array}{l} Bx = 0 \Longrightarrow xAx > 0 \\ x \neq 0 \end{array} \right\rangle \Longleftrightarrow \left\langle \begin{array}{l} A + \alpha B^T B \text{ is positive} \\ \text{definite for all } \alpha \geqq \bar{\alpha} \\ \text{for some } \bar{\alpha} > 0 \end{array} \right\rangle$$

Combining this lemma with Corollary 3.5 gives

Theorem 3.5 (Geometric characterization of positive definiteness of a nonsingular matrix) The nonsingular matrix A in $R^{n\times n}$ is positive definite if and only if it is positive definite on some or each subspace of R^n

$$S = \{x \mid Bx = 0\}, \quad \text{where} \quad B \in R^{m\times n},$$

and A^{-1} is positive semidefinite on the orthogonal complement $S^{\perp}$ of S:

$$S^{\perp} = \{y \mid y = B^T u\} .$$

We remark here that the matrix A may not even be positive semidefinite when A is positive definite on both the space S and its orthogonal complement $S^{\perp}$. This can be seen from the example: $A := \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}$, $S := \{x \in R^2 \mid x_2 = 0\}$.

We are now ready to present our characterization results.

Theorem 3.6 (Positive definiteness of Hessian of Lagrangian under primal second order sufficiency (necessity) and dual second order necessity (sufficiency)) Let $(\bar{x},\bar{u},\bar{v})$ be a Karush-Kuhn-Tucker triple of the primal program (P) that satisfies the strict complementarity condition: $\bar{u}_i > 0$ whenever $g_i(\bar{x}) = 0$, and let the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ be nonsingular. If $(\bar{x},\bar{u},\bar{v})$ satisfies the primal second order sufficient (necessary) optimality condition and the dual second order necessary (sufficient) optimality condition, then $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ is positive definite.

Proof Notice that when the strict complementarity condition holds K is empty and the tangent cone T defined in Section 2 is a subspace and the normal cone N is its orthogonal complement. When the primal second order sufficient (necessary) condition is satisfied the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ is positive definite (semidefinite) on T. While, when the dual second order necessary (sufficient) condition holds, the inverse Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})^{-1}$ is positive semidefinite (definite) on N. Therefore, it follows from Theorem 3.5 that the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ must be positive definite. $\square$

In light of the above theorem, it is natural to expect that the Hessian would be positive semidefinite when both the primal and the dual second order necessary conditions hold. Curiously this turns out not to be true as can be seen from the following example:

$$\begin{array}{ll} \min & -x_1 x_2 \\ \text{s.t.} & x_1 = 0 \end{array}$$

The vector $\bar{x}^T = (0,1)$ together with $\bar{v} = 1$ constitute a Karush-Kuhn-Tucker point. Both the Hessian of the Lagrangian and its inverse are $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$. We also have the tangent space $T = \{x \in R^2 | x_1 = 0\}$ and the normal space $N = \{x \in R^2 | x_2 = 0\}$. Therefore, both the primal and the dual second order necessary conditions hold. But the matrix $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ is not positive semidefinite.

The positive definiteness of the Hessian $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ clearly implies the primal second order sufficient condition. By Theorem 2.3

under the constraint qualification (2.10), the positive definitness of $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ also implies the dual second order sufficient optimality condition. Therefore, we have a converse to Theorem 3.6 which extends and sharpens Luenberger's local duality result [5].

Theorem 3.7 (Primal and dual second order sufficiency under positive definiteness of Hessian of Lagrangian and constraint qualification) Let $\bar{x}$ be a local minimum point of (P) satisfying the constraint qualification (2.10), let f, g and h be twice continuously differentiable at $\bar{x}$ and let $(\bar{u},\bar{v})$ be a Lagrange multiplier associated with $\bar{x}$. If $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})$ is positive definite then $(\bar{x},\bar{u},\bar{v})$ satisfies both the primal and the dual second order sufficient optimality conditions.

References

1. G. Debreu: "Definite and semidefinite quadratic forms", Econometrica 20, 1952, 295-300.

2. A. V. Fiacco & G. P. McCormick: "Nonlinear programming: Sequential unconstrained minimization techniques", Wiley, New York, 1968.

3. P. Finsler: "Über das Vorkommen definiter und semidefinitir Formen in Scharen quadratischen Formen", Commentarii Mathematicii Helveticii 9, 1937, 188-192.

4. S.-P. Han & O. L. Mangasarian: "Exact penalty functions in nonlinear programming", Mathematical Programming 17, 1979, 251-269.

5. D. G. Luenberger: "Introduction to linear and nonlinear programming", Addison-Wesley, Menlo Park, California, 1973.

6. G. P. McCormick: "Second order conditions for constrained minima", SIAM Journal on Applied Mathematics 15, 1967, 641-652.

7. O. L. Mangasarian: "Nonlinear programming", McGraw-Hill, New York, 1969.

8. O. L. Mangasarian: "Unconstrained Lagrangians in nonlinear programming", SIAM Journal on Control 13, 1975, 772-791.

9. O. L. Mangasarian & S. F. Fromvitz: "The Fritz-John necessary optimality conditions in the presence of equality constraints", Journal of Mathematical Analysis and Applications 17, 1967, 34-47.

10. D. H. Martin and D. H. Jacobson: "Copositive matrices and definiteness of quadratic forms subject to homogeneous linear inequality constraints", Linear Algebra and Its Applications 35, 1981, 227-258.

11. J. M. Ortega and W. C. Rheinboldt: "Iterative solution of nonlinear equations in several variables", Academic Press, New York, 1970.

12. T. Pietrzykowski: "An exact potential method for constrained maxima", SIAM Journal on Numerical Analysis 6, 1969, 299-304.

13. S. M. Robinson: "Generalized equations and their solutions; Part II: Applications to nonlinear programming", Technical Summary Report No. 2048, Mathematics Research Center, University of Wisconsin-Madison, 1980, to appear in Mathematical Programming Studies.

14. P. Wolfe: "A duality theorem for nonlinear programming", Quarterly of Applied Mathematics 19, 1961, 239-244.