JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 31, No. 2, JUNE 1980

Second-Order and Related Extremality Conditions in Nonlinear Programming[1]

A. BEN-TAL[2]

Communicated by M. Avriel

Abstract. This paper is concerned with the problem of characterizing a local minimum of a mathematical programming problem with equality and inequality constraints. The main object is to derive second-order conditions, involving the Hessians of the functions, or related results where some other curvature information is used. The necessary conditions are of the Fritz John type and do not require a constraint qualification. Both the necessary conditions and the sufficient conditions are given in equivalent pairs of primal and dual formulations.

Key Words. Nonlinear programming, local extrema second-order conditions, constraint qualification, extremality conditions.

1. Introduction

In this paper, we derive local optimality conditions for the following nonlinear programming problem [Problem (NLP)]:

$$
\begin{aligned}
\text{(NLP)}\quad & \min f^{0}(x),\\
& \text{subject to } f^{k}(x) \leq 0, \qquad k \in I \triangleq \{1, 2, \ldots, P\},\\
& \qquad h^{j}(x) = 0, \qquad j \in J \triangleq \{1, 2, \ldots, m\},\\
& \qquad x \in R^{n}.
\end{aligned}
$$

No special assumptions, other than differentiability, are imposed on the problem's functions. In particular, a constraint qualification is not needed for the validity of our necessary condition. The main object is to derive

[1] This research was partly supported by Project No. NR-947-021, ONR Contract No. N00014-75-0569, with the Center for Cybernetic Studies, and by the National Science Foundation, Grant No. NSF-ENG-76-10260.

[2] Associate Professor, Department of Computer Science, Technion-Israel Institute of Technology, Haifa, Israel.

0022-3239/80/0600-0143$03.00/0

Table 1. Second-order conditions for Problem (NLP).

	Primal form	Dual form
Necessary conditions for local minimum	For every $d \in D(x^*)$, there is no z such that $(z, d) \neq 0$ and $\nabla f^k(x^*)z + d'\nabla^2 f^k(x^*)d < 0,\quad k \in \Lambda^*(d)$, $\nabla h^j(x^*)z + d'\nabla^2 h^j(x^*)d = 0,\quad j \in J$. Here, it is assumed that $\{\nabla h^j(x^*)\}_{j\in J}$ are linearly independent.	For every $d \in D(x^*)$, there exist nonnegative $y \in R^{P+1}$, $\mu \in R^m$, $(y, \mu) \neq 0$, such that $\nabla L(x^*, y, \mu) = 0$, $d'\nabla^2 L(x^*, y, \mu)d \geq 0$, $y_k f^k(x^*) = 0,\quad k \in I^*$, $y_k \nabla f^k(x^*)d = 0,\quad k \in I_0^*$.
Sufficient conditions for (isolated) local minimum	For every $d \in D(x^*)$, there is no z such that $(z, d) \neq 0$ and $\nabla f^k(x^*)z + d'\nabla^2 f^k(x^*)d \leq 0,\quad k \in \Lambda^*(d)$, $\nabla h^j(x^*)z + d'\nabla^2 h^j(x^*)d = 0,\quad j \in J$.	For every nonzero $d \in D(x^*)$, there exist nonnegative $y \in R^{p+1}$, $\mu \in R^m$. $(y, \mu) \neq 0$, such that $\nabla L(x^*, y, \mu) = 0$, $d'\nabla^2 L(x^*, y, \mu)d > 0$, $y_k f^k(x^*) = 0,\quad k \in I^*$, $y_k \nabla f^k(x^*)d = 0,\quad k \in I_0^*$.

second-order conditions, involving the Hessians of the problem functions, or derive related results where some other curvature information is used (see, e.g., Corollaries 2.1, 2.3, 2.4).

To illustrate the distinguished character of the second-order conditions obtained here, and to discuss their relations to the classical results, we collect them in Table 1. The following functions and sets are used in this table and throughout the paper:

(i) the *Lagrangian function* $L: R^n \times R^{P+1} \times R^m \to R$,

$$L(x, y, \mu) \triangleq \sum_{k\in\{0\}\cup I} y_k f^k(x) + \sum_{j\in J} \mu_j h^j(x);$$

(ii) the *set of critical directions* at x^*,

$$D(x^*) \triangleq \{d \in R^n : \nabla f^k(x)d \leq 0,\ k \in I_0^*\ ;\ \nabla h^j(x^*)d = 0,\ j \in J\},$$

where

$$I_0^* \triangleq \{0\} \cup I^*, \qquad I^* \triangleq \{k \in I : f^k(x^*) = 0\},$$

and

$$\Lambda^*(d) \triangleq \{k \in I_0^* : \nabla f^k(x^*)d = 0\}.$$

We write $\nabla L(x^*, y, \mu)$ for the gradient of $L(\cdot, y, \mu)$ evaluated at $x = x^*$. Gradient vectors are always row vectors. All other vectors (such as x, d, etc.) are column vectors. Row vectors will be denoted by primes, e.g., x', d', etc.

The following conclusions may be drawn from Table 1.

(i) The dual necessary conditions include the first-order result of Mangasarian and Fromovitz (Ref. 1) and in particular (for problem without equality constraints, i.e., $J = \emptyset$) the Fritz John conditions (Ref. 2).

(ii) For an unconstrained problem, i.e., $I = \emptyset$, $J = \emptyset$, both the primal and the dual conditions reduce to the well-known sufficient condition [$\nabla f^o(x^*) = 0$, $\nabla^2 f^o(x^*)$ positive definite] and necessary condition [$\nabla f^o(x^*) = 0$, $\nabla^2 f^o(x^*)$ positive semidefinite].

(iii) For problems with equality constraints only ($I = \emptyset$), the dual results reduce to the necessary and the sufficient conditions given by McShane (Ref. 3).

The reader will notice that the dual conditions do not speak about the existence of *fixed multipliers*, but rather about the existence of multipliers which are *functions of the critical directions* $d \in D(x^*)$. This is in contrast to the Kuhn–Tucker type second-order necessary conditions (e.g., Ref. 4) or the commonly used sufficient conditions of Penissi (Ref. 5) or McCormick (Ref. 6); see also Ref. 4. The latter are special cases of our dual sufficiency result. Under certain constraint qualifications, the dual necessary conditions reduce to some classical Kuhn–Tucker type conditions, e.g., Theorem 3.2. Moreover, new types of constraint qualifications can now be formulated (see Proposition 3.1).

Historically, optimality conditions were derived first for *equality constrained* problems. It was a common practice in later years to handle *inequalities* by considering only *active constraints* and treat them as equalities (e.g., Refs. 7–9). This approach, however, cannot produce the type of results given in Table 10. See Example 3.2 and Remark 3.1. Our approach is to treat programs with *inequalities only*, first. Programs with equalities are then treated by solving the equations, eliminating part of the variables, and reducing the original problem to one with fewer variables and with inequalities only. This is essentially the approach taken by Mangasarian and Fromovitz (Ref. 1) in extending the Fritz John condition to problems with equality constraints. They also observed that the naive approach of writing an equation $h(x) = 0$ as two inequalities [$h(x) \leq 0$, $-h(x) \leq 0$] is useless as far as *necessary conditions* are concerned, because the resulting conditions are trivially satisfied at *any feasible point*. The same is true for second-order conditions; see Section 3.

2. Necessary Conditions for Problems with Inequalities Only

In this section, we study a special case of Problem (NLP), where $J=\varnothing$. Thus, we consider the following problem [Problem (P)]:

(P) $\min f^0(x)$,

subject to $f^k(x)\leq 0, \qquad k\in I.$

First, we recall the following definition: a differentiable function $g: R^n\to R$ is *pseudoconcave* on the convex subset S at x if

$$\nabla g(x)(z-x)\leq 0 \Rightarrow g(z)\leq g(x), \qquad \text{for all } z\in S.$$

The function g is *strictly pseudoconcave* if

$$\nabla g(x)(z-x)\leq 0 \Rightarrow g(z)< g(x), \qquad z\neq x.$$

In particular, a function $\gamma: R\to R$ is pseudoconcave at 0^+ if, for some $T>0$, $\gamma(t)$ is pseudoconcave on $[0, T]$ at $t=0$. For a fixed vector $x^*\in R^n$ and $d\in R^n$, let

$Q_k^*(d)\triangleq\{z: (z,d)\neq 0, f^k(x^*+td+\frac{1}{2}t^2z)$ is pseudoconcave at 0^+ as a function of $t\}$, $k\neq 0$;

$Q_0^*(d)\triangleq\{z: (z,d)\neq 0, f^0(x^*+td+\frac{1}{2}t^2z)$ is strictly pseudoconcave at 0^+ as a function of $t\}$.

The following lemma is a convenient tool to derive a variety of primal necessary conditions.

Lemma 2.1. Let the functions $\{f^k: k\in\{0\}\cup I\}$ be differentiable, and let x^* be a local minimizer of Problem (P). Then, for every d satisfying

$$\nabla f^k(x^*)d\leq 0, \qquad k\in I_0^*, \tag{1}$$

it follows that $Q^*(d)=\varnothing$, where

$$Q^*(d)\triangleq \bigcap_{k\in\Lambda^*(d)} Q_k^*(d).$$

Proof. Let $\bar{d}$ be a vector satisfying (1), and assume that $Q^*(\bar{d})\neq\varnothing$. Let then $\bar{z}\in Q^*(\bar{d})$. For $k\in I\backslash I^*$, $f^k(x^*)<0$; hence, by continuity, for some $T_k>0$ sufficiently small,

$$f^k(x^*+t\bar{d}+\tfrac{1}{2}t^2\bar{z})\leq 0, \qquad t\in[0,T_k],\ k\in I\backslash I_0^*. \tag{2}$$

For every $k\in I_0^*\backslash\Lambda^*(\bar{d})$, we have that $\nabla f^k(x^*)\bar{d}<0$, but

$$0>\nabla f^k(x^*)\bar{d}=[(d/dt)f^k(x^*+t\bar{d}+\tfrac{1}{2}t^2\bar{z})]_{t=0},$$

showing that $f^k(x^*+t\bar{d}+\frac{1}{2}t^2\bar{z})$ is strictly decreasing; hence, for some $T_k>0$,

$$f^k(x^*+t\bar{d}+\tfrac{1}{2}t^2\bar{z})<0, \qquad t\in[0,T_k], \qquad k\in I^*\backslash\Lambda^*(\bar{d}); \tag{3}$$

and, if $0\notin\Lambda^*(\bar{d})$, also

$$f^\circ(x^*+t\bar{d}+\tfrac{1}{2}t^2\bar{z})<f^\circ(x^*), \qquad t\in[0,T_0], \quad \text{for some } T_0>0. \tag{4}$$

For every $k\in\Lambda^*(\bar{d})$, we have $\nabla f^k(x^*)\bar{d}=0$; hence, since $\bar{z}\in Q^*(\bar{d})$, for some $T_k>0$,

$$f^k(x^*+t\bar{d}+\tfrac{1}{2}t^2\bar{z})\leq f^k(x^*)=0, \qquad t\in[0,T_k], k\in\Lambda^*(\bar{d}), k\neq 0;$$

and, if $0\in\Lambda^*(\bar{d})$ also,

$$f^\circ(x^*+t\bar{d}+\tfrac{1}{2}t^2\bar{z})<f^\circ(x^*), \qquad t\in[0,T_0], \quad \text{for some } T_0>0. \tag{5}$$

One infers from (2)–(5) that a small movement from x^* along the curve

$$\bar{e}(t)=x^*+t\bar{d}+\tfrac{1}{2}t^2\bar{z}$$

decreases the value of the objective function at x^*, while maintaining feasibility. This of course violates the assumption that x^* is a local minimizer. □

Lemma 2.1 is a source to derive both first-order and second-order optimality conditions. First, we derive the following refinement of the Fritz John condition obtained by Mangasarian (Ref. 10, Theorem 10.2.2). We recall that a function h is *locally pseudoconcave* at x if, for some neighborhood N of x, h is pseudoconcave on N at x. Let us define also the following index set:

$$N^*\triangleq\{k\in I: f^k \text{ is not locally pseudoconcave at } x^*\}.$$

Corollary 2.1. Under the assumptions of Lemma 2.1, a necessary condition for a feasible point x^* to be a local minimizer of Problem (P) is that the system [System (I)]

$$\begin{aligned}&\nabla f^\circ(x^*)d<0,\\ &\nabla f^k(x^*)d<0, \qquad k\in N^*,\\ &\nabla f^k(x^*)d\leq 0, \qquad k\in I^*\backslash N^*,\end{aligned}$$

has no solution $d\in R^n$, or equivalently that the system [System (II)]

$$\begin{aligned}&\sum_{k\in I_0^*} y_k\nabla f^k(x^*)=0,\\ &y\geq 0, \qquad k\in I_0^*,\\ &\{y_k: k\in\{0\}\cup N^*\} \text{ are not all zero}\end{aligned}$$

has a solution $\{y_k: k\in I_0^*\}$.

Proof. Only the inconsistency of System (I) has to be shown, since the equivalent statement, on the consistency of System (II), follows from Motzkin's theorem of the alternatives (e.g., Ref. 10). Note that, for $k \notin N^*$, the corresponding function f^k is locally pseudoconcave at x^*; hence, in particular, $f^k(x+td)$ is pseudoconcave at 0^+, as a function of t, for every $d \in R^n$. This implies that

$$0 \in \bigcap_{k \notin N^*} Q_k^*(d), \qquad \text{for every } d.$$

If System (I) has a solution, say $\bar{d}$, then

$$\Lambda^*(\bar{d}) \subset I^* \backslash N^*;$$

and so, from the previous relation,

$$0 \in \bigcap_{\Lambda^*(d)} Q_k^*(\bar{d}) \neq \varnothing,$$

violating the necessary condition of Lemma 2.1. □

Let $\{N_k^*(d)\}$ be subsets satisfying

$$N_k^*(d) \subset Q_k^*(d), \qquad k \in \Lambda^*(d). \tag{6}$$

Then, the condition $Q^*(d) = \varnothing$ clearly implies that

$$\bigcap_{k \in \Lambda^*(d)} N_k^*(d) = \varnothing;$$

hence, replacing the set $Q^*(d)$ by $\cap N_k^*(d)$ in Lemma 2.1 results in a valid necessary condition. This idea can be exploited in various ways. Perhaps, the most natural way is the one used to obtain the following primal necessary condition.

Corollary 2.2. Let $\{f^k: k \in \{0\} \cup I\}$ be twice continuously differentiable, and suppose that x^* is a local minimizer of Problem (P). Then, for every d satisfying (1), it follows that no $z \in R^n$ solves the system

$$\nabla f^k(x^*)z + d' \nabla^2 f^k(x^*) d < 0, \qquad k \in \Lambda^*(d). \tag{7}$$

Proof. Note that (7) is the same as

$$[(d^2/dt^2) f^k(x^* + td + \tfrac{1}{2}t^2 z)]_{t=0} < 0;$$

the latter implies that $f^k(x^* + td + \frac{1}{2}t^2 z)$ is strictly concave in the neighborhood of $t=0$; hence, it is strictly pseudoconcave at 0^+. Therefore, the set $N_k^*(d)$ of all z's solving (7) satisfies (6). □

The dual necessary conditions follow.

Theorem 2.1. Let $\{f^k: k \in \{0\} \cup I\}$ be twice continuously differentiable functions, and suppose that x^* is a local minimizer of Problem (P). Then, corresponding to every vector d satisfying (1), there exist multipliers

$$y_k \geqslant 0, \qquad k \in I_0^*, \text{ not all zero,} \tag{8}$$

satisfying

$$\sum_{k \in I_0^*} y_k \nabla f^k(x^*) = 0, \tag{9}$$

$$y_k \nabla f^k(x^*) d = 0, \qquad k \in I_0^*, \tag{10}$$

$$d' \Big[\sum_{k \in I_0^*} y_k \nabla^2 f^k(x^*) \Big] d \geqslant 0. \tag{11}$$

Proof. Let $\bar{d}$ be a fixed but arbitrary vector satisfying (1). Consider the matrix A whose rows are

$$\{\nabla f^k(x^*) : k \in \Lambda^*(\bar{d})\},$$

and consider the vector b whose components are

$$\{-\bar{d}' \nabla^2 f^k(x^*) \bar{d} : k \in \Lambda^*(\bar{d})\}.$$

With these notations, Corollary 2.2 states that the linear system

$$\{Az < b\}$$

has no solution. This is equivalent to saying that the linear program

$$\max(\lambda : Az + \vec{\lambda} \leqslant b\}$$

has optimal value $\lambda^* \leqslant 0$; here, $\vec{\lambda}$ is the vector $(\lambda, \lambda, \ldots, \lambda)'$, with obvious dimensionality. Thus, the dual program

$$\min\{b'y : A'y = 0, \Sigma y_k = 1, y_k \geqslant 0\}$$

has a nonpositive optimal value, i.e., the system

$$A'y = 0, \qquad b'y \leqslant 0, \qquad y \geqslant 0, \qquad y \neq 0, \tag{12}$$

has a solution

$$\bar{y} = \{\bar{y}_k : k \in \Lambda^*(\bar{d})\}.$$

If we define also $\bar{y}_k = 0$ for every $k \in I_0^* \backslash \Lambda^*(\bar{d})$, we see from (12) that

$$\{\bar{y}_k : k \in I_0^*\}$$

satisfy (8)–(11). □

It should be emphasized that the multipliers are not *fixed constants*, but rather *functions of the critical directions* [i.e., those d's which satisfy (1)]. In the following example, we illustrate a situation where, at an optimal point, there are no fixed multipliers satisfying (8)–(11). This means, in particular, that none of the Kuhn–Tucker type second-order conditions are valid for this example.

Example 2.1. Consider the problem

$$\min f^\circ = 2x_1x_2 + \tfrac{1}{2}x_3^2,$$
$$\text{subject to } f^1 = 2x_1x_3 + \tfrac{1}{2}x_2^2 \leqslant 0,$$
$$f^2 = 2x_2x_3 + \tfrac{1}{2}x_1^2 \leqslant 0.$$

The point

$$x^* = (0, 0, 0)'$$

is an isolated local (in fact, global) minimizer. This can be easily verified as follows. If there is an $x \neq 0$ at least as good as $x^* = 0$, then it must satisfy $f^\circ(x) \leqslant 0$, in addition to

$$f^1(x) \leqslant 0, \qquad f^2(x) \leqslant 0.$$

If one of the components of x is zero, say $x_1 = 0$, then it follows, from

$$f^1 \leqslant 0, \qquad f^2 \leqslant 0,$$

that $x_2 = x_3 = 0$, hence $x = 0$. Thus, $x_1 \neq 0$, $x_2 \neq 0$, $x_3 \neq 0$; this implies that

$$x_1x_2 < 0, \qquad x_1x_3 < 0, \qquad x_2x_3 < 0,$$

which is impossible, since multiplication of these three inequalities yields

$$x_1^2x_2^2x_3^2 < 0.$$

Note that here

$$\nabla f^k(x^*) = 0, \qquad i = 0, 1, 2;$$

hence, (1) is satisfied by *every* $d \in R^n$ and (9), (10) are satisfied by arbitrary multipliers. Thus, only (8) and (11) need checking. We show now that no *fixed multipliers* satisfying (8) can satisfy condition (11), which is here

$$(d_1, d_2, d_3)\begin{bmatrix} y_2 & 2y_0 & 2y_1 \\ 2y_0 & y_1 & 2y_2 \\ 2y_1 & 2y_2 & y_0 \end{bmatrix}\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \geqslant 0, \qquad d \in R^n. \tag{13}$$

Indeed, for

$$d^A = (0, 1, -1)', \qquad d^B = (1, 0, -1)', \qquad d^C = (-1, 1, 0)',$$

(13) become, respectively,

$$y_0+y_1-4y_2\geqslant 0,$$
$$y_0+y_2-4y_1\geqslant 0,$$
$$y_1+y_2-4y_0\geqslant 0,$$

which upon adding yield

$$2(y_0+y_1+y_2)\leqslant 0,$$

contradicting (8).

However, (8) and (11) do hold with

$$(y_0, y_1, y_2)=\begin{cases}(1, 0, 0), & \text{if } d_1d_2\geqslant 0,\\ (0, 1, 0), & \text{if } d_1d_3\geqslant 0,\\ (0, 0, 1), & \text{if } d_2d_3\geqslant 0,\end{cases}$$

showing that Theorem 2.1 is valid.

Another consequence of Lemma 2.1 is the following first-order result which, in certain special cases (to be discussed below), may do better than the second-order results.

Corollary 2.3. Under the assumptions of Lemma 2.1, a necessary condition for a feasible point x^* to be a local minimizer of Problem (P) is that, for every $\bar{d}\neq 0$ satisfying (1),

$$\bar{d}\notin \bigcap_{k\in\Lambda^*(\bar{d})} L_k^*,$$

where

$$L_k^*\triangleq\begin{cases}\{d: f^k(x^*+td) \text{ is concave at } 0^+\}, & k>0,\\ \{d: f^\circ(x+td) \text{ is strictly concave at } 0^+\}, & k=0.\end{cases}$$

Proof. Suppose that $\bar{d}$ is a nonzero vector satisfying (1) and also $\bar{d}\in\cap L_k^*$. Then, $0\in Q^*(\bar{d})$, violating the necessary conditions of Lemma 2.1. □

To illustrate this last result, consider the following example.

Example 2.2.

$$\min f^\circ = x_1,$$
$$\text{subject to } f^1=-(1-x_1)^3+x_2\leqslant 0,$$
$$f^2=-x_1\leqslant 0,$$
$$f^3=-x_2\leqslant 0.$$

This is the famous example used to demonstrate the failure of the Kuhn–Tucker conditions in the lack of constraint qualification. The second-order condition in Theorem 2.1 does not help much, because these conditions are trivially satisfied here, regardless of the objective function. In particular, they are satisfied at $x^* = (1, 0)$, which is not a local minimum here (in fact, a global maximum). At this point, the active constraints are the first and the third, so

$$I_0^* = \{0, 1, 3\}.$$

In applying Corollary 2.3, note that f^0 and f^3 are concave and that the inequalities

$$\nabla f^1(x^*)d < 0, \qquad \nabla f^3(x^*)d < 0$$

are contradictory. Thus, we are left to check whether there exists $(d_1, d_2) \neq 0$, such that

$$d_1 \leq 0, \qquad d_2 = 0, \qquad t^3 d_1^3 \text{ is concave at } 0^+, \qquad -d_2 \leq 0; \tag{14}$$

since

$$(d_1, d_2) \neq 0, \qquad d_2 = 0,$$

it follows from the first constraint that $d_1 < 0$; hence, $t^3 d_1^3$ is strictly concave for $t > 0$, and hence the system (14) is consistent, showing that x^* is indeed nonoptimal.

Note that, if f^k is a convex function, then, for every $d \in R^n$,

$$L_k^* = \begin{cases} \{d : f^k(x^* + td) \text{ is linear at } 0^+\}, & k > 0, \\ \varnothing, & k = 0. \end{cases} \tag{15}$$

Moreover, for $0 \neq k \in \Lambda^*(d)$,

$$L_k^* = \{d : f^k(x^* + td) \text{ is constant at } 0^+\} \triangleq D_k^*. \tag{16}$$

The set D_k^* was introduced in Ref. 11, where it was called the *cone of directions of constancy* of f^k at x^*. It was shown there that this cone is also convex and is in general quite easy to compute. In fact, for analytic convex functions, it can be described by a system of homogeneous linear equations.

Thus, for a convex programming problem, Corollary 2.3 reduces to the following result, which was first obtained in Ref. 11.

Corollary 2.4. Let $\{f^k : k \in \{0\} \cup I\}$ be differentiable convex functions, and let x^* be a minimum point of Problem (P). Then, for every d satisfying

$$\nabla f^0(x^*)d < 0,$$
$$\nabla f^k(x^*)d \leq 0, \qquad k \in I^*,$$

it follows that

$$d \notin \bigcap_{\Lambda^*(d)} D_k^*.$$

Remark 2.1. The optimality conditions in Corollary 2.4 are also *sufficient*, the reason being that, for a convex feasible set, it is enough to check improvement of a given solution only along straight lines.

Another important case where Corollary 2.4 gives a *necessary and sufficient criterion* for local minimum is when the feasible set F of (P) is *locally starshaped* at x^*; i.e., there exists a neighborhood N of x^*, with $N \cap F \neq \emptyset$, such that

$$x \in N \cap F, \qquad 0 \leq \lambda \leq 1 \Rightarrow (1-\lambda)x^* + \lambda x \in N \cap F.$$

F is called *locally starshaped* if it is locally starshaped at every $x \in F$. Any set which can be represented as a finite union of convex sets has this property. Examples of functions f for which the constraint $f \leq 0$ generates a locally starshaped set are

$$f(x) = \prod_{i=1}^{k} (a_i'x + \beta), \qquad f(x) = -(a'x + \beta)^2 + \phi^2(x);$$

here, ϕ is a convex function.

3. Necessary Conditions for Problem (NLP)

In principle, an equality constraint $h^j(x) = 0$ can be expressed as two inequalities:

$$h^j(x) \leq 0, \qquad -h^j(x) \leq 0. \tag{17}$$

However, this does not help much as far as the *necessary conditions* are concerned, because the conditions in Theorem 2.1 will be trivially satisfied at an *arbitrary feasible solution*. In the primal version of Theorem 2.1 (Corollary 2.2), we will have the following pair of contradicting inequalities, which result from (7) when applied to (17):

$$\nabla h^j(x^*)z + d'\nabla^2 h^j(x^*)d < 0,$$
$$-\nabla h^j(x^*)z - d'\nabla^2 h^j(x^*)d < 0.$$

The approach adopted here instead is to eliminate the equality constraints by solving the equations

$$\{h^j(x) = 0, j \in J\},$$

thus expressing some of the variables in terms of the others. This reduces Problem (NLP) to Problem (P) with fewer variables and with inequalities only. The gradients and Hessians of the reduced functions are computed by using the implicit function theorem.

First, we extend Corollary 2.2.

Theorem 3.1. *Primal Necessary Condition.* Consider Problem (NLP) with twice continuously differentiable functions

$$\{f^k : k \in \{0\} \cup I\}, \qquad \{h^j : j \in J\},$$

and suppose that x^* is a local minimizer of Problem (NLP). Assume further that the gradients $\{\nabla h^j(x^*): j \in J\}$ are linearly independent. Then, for every critical direction d, i.e., for every d satisfying

$$\nabla f^k(x^*)d \leq 0, \qquad k \in I_0^*, \tag{18}$$

$$\nabla h^j(x^*)d = 0, \qquad j \in J, \tag{19}$$

it follows that no $z \in R^n$ solves the system

$$\nabla f^k(x^*)z + d'\nabla^2 f^k(x^*)d < 0, \qquad k \in \Lambda^*(d), \tag{20}$$

$$\nabla h^j(x^*)z + d'\nabla^2 h^j(x^*)d = 0, \qquad j \in J. \tag{21}$$

Proof. If $m = \text{card}\, J$ is equal to n (number of variables), then $d = 0$ is the only solution of (19), and so $z = 0$ is the only solution of (21); hence, (20) cannot hold. Thus, for this case, the optimality conditions are trivially satisfied. Therefore, we assume that $m < n$; and so, by the assumption that $\{\nabla h^j(x^*): j \in J\}$ are linearly independent, it follows that there exists a partition

$$L \cup B = \{i, 2, \ldots, n\}, \qquad \text{card}\, B = m,$$

such that the matrix H whose (i, j)th element is

$$H_{ij} = \left\{\frac{\partial h^j}{\partial x_i}(x^*): j \in J, i \in B\right\}$$

is nonsingular. Therefore, by the implicit function theorem, applied to the system

$$\{h^j(x) = 0, h \in J\},$$

we conclude that, for some neighborhood N of x^*, there exists a twice continuously differentiable vector-valued function $\theta: R^{n-m} \to R^m$, with components $\{\theta^r: r \in B\}$, such that

$$x_B = \theta(x_L), \qquad \text{for all } x = (x_L, x_B) \in N,$$

i.e.,

$$h^j(x_L, \theta(x_L)) \equiv 0, \qquad j \in J. \tag{22}$$

Problem (NLP) is then reduced to the following inequality constrained problem with $n-m$ variables [Problem $(\hat{\mathrm{P}})$]:

$(\hat{\mathrm{P}})$ $\min\{\hat{f}^\circ(x_L)$: subject to $\hat{f}^k(x_L) \leq 0$, $k \in I\}$, where

$$\hat{f}(x_L) \triangleq f^k(x_L, \theta(x_L)), \qquad k \in \{0\} \cup I, \tag{23}$$

and our assumption implies that x_L^* is a local minimum for Problem $(\hat{\mathrm{P}})$.

Suppose now that the conclusion of the theorem is false, and let d, z be a pair of vectors satisfying (18)–(21). To proceed, we first introduce some notations: $\nabla_L h^J$ is the $m \times (n-m)$ matrix whose jth row, $j \in J$, has components

$$\left\{\frac{\partial h^j}{\partial x_i}(x^*): i \in L\right\}.$$

Likewise, $\nabla_B h^s$ is the row vector with components

$$\left\{\frac{\partial h^s}{\partial x_i}: i \in B\right\}.$$

Also,

$$\theta_j^k \triangleq \frac{\partial \theta^k}{\partial x_j}(x_L^*), \qquad \theta_{ij}^k \triangleq \frac{\partial^2 \theta^k}{\partial x_i \partial x_j}(x_L^*);$$

h_j^k and h_{ij}^k have similar meanings.

With the above notations, (19) can be rewritten as

$$H d_B + \nabla_L h^J d_L = 0,$$

from which

$$d_B = -H^{-1}(\nabla_L h^J) d_L. \tag{24}$$

Implicit differentiation of (22) gives

$$\nabla \theta = -H^{-1}(\nabla_L h^J),$$

so

$$\begin{aligned} \nabla \hat{f}^k &= \nabla_L f^k + (\nabla_B f^k) \nabla \theta, && \text{by (23)}, \\ \nabla \hat{f}^k &= \nabla_L f^k - (\nabla_B f^k) H^{-1}(\nabla_L h^J), && \text{by (24)}; \end{aligned} \tag{25}$$

hence, by (18) and (24),

$$(\nabla_L \hat{f}^k) d_L \leq 0, \qquad k \in I_0^*. \tag{26}$$

Rewriting (21) as

$$Hz_B + \nabla_L h^J z_L = -w,$$

where w is a vector whose rth component is

$$w_r \triangleq d'\nabla^2 h^r(x^*)d, \tag{27}$$

one infers that

$$z_B = -H^{-1}w - H^{-1}(\nabla_L h^J)z_L. \tag{28}$$

Twice differentiation of (22) gives

$$h^s_{ij} + \sum_{k\in B} h^s_{ik}\theta^k_j + \sum_{t\in B} h^s_{tj}\theta^t_i + \sum_{r,m\in B} h^s_{rm}\theta^m_j\theta^r_i + \sum_{r\in B} h^s_r\theta^r_{ij} = 0, \qquad i, j \in L, s \in J.$$

Multiplying these equations by $d_i d_j$, summing over all $i, j \in L$, using the fact that

$$d_k = \sum_{j\in L} d_j\theta^k_j, \qquad k \in B,$$

which follows from (24) and (25), one obtains

$$(\nabla_B h^s)a = -d'(\nabla^2 h^s)d, \qquad s \in J, \tag{29}$$

where a is the vector rth component is

$$a_r \triangleq d'_L(\nabla^2\theta^r)d_L. \tag{30}$$

Similar to the derivation of (29), we derive from (23)

$$d'_L(\nabla^2_{LI}\hat{f}^k)d_L = d'(\nabla^2 f^k)d + (\nabla_B f^k)a. \tag{31}$$

But, from (29) and (27),

$$a = -H^{-1}w = z_B + H^{-1}(\nabla_L h^J)z_L,$$

by (28), so that (31) becomes

$$\begin{aligned} d'_L(\nabla^2_L\hat{f}^k)d_L &= d'(\nabla^2 f^k)d + (\nabla_B f^k)z_B + (\nabla_B f^k)H^{-1}(\nabla_L h^J)z_L \\ &= d'(\nabla^2 f^k)d + (\nabla_B f^k)z_B + (\nabla_L f^k)z_L - (\nabla_L\hat{f}^k)z_L, \end{aligned}$$

by (25); or, rearranging terms,

$$(\nabla_L\hat{f}^k)z_L + d'_L(\nabla^2_L\hat{f}^k)d_L = d'(\nabla^2 f^k)d + (\nabla f^k)z < 0, \qquad \text{for } k \in \Lambda^*(d),$$

by (20). The latter inequalities, together with Ineqs. (26), show that $d_L, z_L \in R^{n-m}$ are vectors violating the necessary conditions for x^*_L to be a local minimum of Problem ($\hat{\mathrm{P}}$), a contradiction. □

The dual result of Theorem 3.1 is an extension to second-order conditions of a result due to Mangasarian and Fromovitz (Ref. 1) and, for $J = \emptyset$, of the classical Fritz John conditions.

Theorem 3.2. *Dual Necessary Conditions.* Consider Problem (NLP) with twice continuously differentiable functions

$$\{f^k: k \in \{0\} \cap I\}, \qquad \{h^j: j \in J\},$$

and suppose that x^* is a local minimizer of Problem (NLP). Then, corresponding to every critical direction d, there exist multipliers

$$y_k \geq 0, \qquad k \in I_0^*, \qquad \mu_j \in R, \qquad j \in J, \text{ not all zero,} \tag{32}$$

satisfying

$$\sum_{k \in I_0^*} y_k \nabla f^k(x^*) + \sum_{j \in J} \mu_j \nabla h^j(x^*) = 0, \tag{33}$$

$$y_k \nabla f^k(x^*) d = 0, \qquad k \in I_0^*, \tag{34}$$

$$d'\left[\sum_{k \in I_0^*} y_k \nabla^2 f^k(x^*) + \sum_{j \in J} \mu_j \nabla^2 h^j(x^*)\right] d \geq 0. \tag{35}$$

Proof. If $\{\nabla h^j(x^*): j \in J\}$ are *linearly dependent*, then (32)–(34) are trivially satisfied by $y_k = 0$, $k \in I_0^*$, and some $\mu_j = \bar{\mu}_j$, $j \in J$, not all zero. Moreover, for every $d \in R^n$, (35) is satisfied either by the above chosen multipliers or by $y_k = 0$, $\mu_j = -\bar{\mu}_j$.

If $\{\nabla h^j(x^*): j \in J\}$ are *linearly independent*, then the assumptions of Theorem 3.1 are satisfied, in which case one can dualize its necessary condition in a manner similar to that used in proving Theorem 2.1. We omit the details. □

It may happen that, for every *critical direction d*, i.e., a direction d satisfying (18) and (19), the corresponding multiplier of the objective function $y_0 = y_0(d)$ is equal to zero, in which event the necessary conditions hold regardless of the objective function. This may happen for example, if x^* is an isolated feasible point. To exclude such possibilities, an additional assumption is needed. Thus, we shall say that the constraints of Problem (NLP) satisfy a *constraint qualification* if, for at least one critical direction d, the multiplier $y_0(d)$ is strictly positive. All the *first-order constraint qualifications* and *second-order constraint qualifications* (see, e.g., Ref. 4, Ref. 6, and Ref. 10), used to obtain the customary Kuhn–Tucker type necessary conditions, fall into the category of our definition, which is broader. See Example 3.1 below. Note that a constraint qualification still does not guarantee the existence of *fixed multipliers* satisfying (32)–(35) for *every* critical direction d. Conditions which *do* guarantee the latter will be termed *strong constraint qualification*. The above mentioned second-order constraint qualifications are strong.

The following Condition (CQ) is a natural extension of the so-called modified Arrow–Hurwitz–Uzawa condition, introduced in Ref. 1; see also Ref. 10, p. 172:

(CQ) The gradients $\{\nabla h^j(x^*): j \in J\}$ are linearly independent and there exists a critical direction d satisfying (18), (19), and a vector $z \in R^n$ such that

$$\begin{aligned} &\nabla f^k(x^*)z + d'\nabla^2 f^k(x^*)d < 0, && k \in \Gamma^*(d),\\ &\nabla h^j(x^*)z + d'\nabla^2 h^j(x^*)d = 0, && j \in J, \end{aligned}$$

where

$$\Gamma^*(d) \triangleq \{k \in I^*: \nabla f^k(x^*)d = 0\}. \tag{36}$$

Note that, if there exist $v \in R^n$ satisfying

$$\begin{aligned} &\nabla f^k(x^*)v < 0, && k \in I^*_{\cdot},\\ &\nabla h^j(x^*)v = 0, && j \in J, \end{aligned}$$

i.e., if the modified Arrow–Hurwitz–Uzawa condition is satisfied, then (CQ) holds trivially with $d = 0$ and $z = v$.

Proposition 3.1. Condition (CQ) is a constraint qualification.

Proof. We have to show that $y_0(d) > 0$ for some critical direction. Assume the contrary, i.e., $y_0 \equiv 0$. Thus, by (32)–(35), the linear system with unknowns $\{y_k: k \in \Gamma^*(d)\}$, $\{\mu_j: j \in J\}$,

$$\begin{aligned} &\sum_{k \in \Gamma^*(d)} y_k \nabla f^k(x^*) + \sum_{j \in J} \mu_j \nabla h^j(x^*) = 0,\\ &\sum_{k \in \Gamma^*(d)} y_k (d'\nabla^2 f^k(x^*)d) + \sum_{j \in J} \mu_j (d'\nabla^2 h^j(x^*)d) \geqslant 0, \end{aligned} \tag{37}$$

$$\begin{aligned} &y_k \geqslant 0, \qquad k \in \Gamma^*(d),\\ &\sum_{k \in \Gamma^*(d)} y_k = 1, \end{aligned} \tag{38}$$

has a solution for every critical direction d. (Note that (38) reflects here condition (32), since the linear independence of $\{\nabla h^j(x^*): h \in J\}$ excludes, by (37), the possibility that $y_k = 0$, for all $k \in \Gamma^*(d)$.)

Therefore, the linear program

$$\min \sum_{k\in\Gamma^*(d)} y_k(-d'\nabla^2 f^k(x^*)d) + \sum_{j\in J} \mu_j(-d'\nabla^2 h^j(x^*)d),$$

subject to (37) and (38),

has a nonpositive optimal value for every critical direction d. Consequently, the dual program

$$\begin{aligned} &\max \lambda, \\ &\text{subject to} \quad \nabla f^k(x^*)z + \lambda \leq -d'\nabla^2 f^k(x^*)d, \qquad k \in \Gamma^*(d), \\ &\qquad\qquad\quad \nabla h^j(x^*)z = -d'\nabla^2 h^j(x^*)d, \qquad j \in J, \end{aligned}$$

has also a nonpositive optimal value $\lambda^* \leq 0$ for every critical direction d, showing that no such d exists for which some $z \in R^n$ satisfies condition (CQ). □

Our next result states essentially that the following condition [Condition (SCQ)]:

(SCQ) $\{\nabla h^j(x^*): j \in J\}, \{\nabla f^k(x^*): k \in I^*\}$ are linearly independent

is a strong constraint qualification. However, we will write it formally, because it is rather a well-worn result encountered in most textbooks on nonlinear programming (e.g., Refs. 4, 7, 8, 9). The proof of this result is quite involved in the above mentioned references; but, with Theorem 3.1 at hand, it is almost trivial.

Theorem 3.3. Under the assumptions of Theorem 3.2 and the additional assumption (SCQ), a necessary condition for a feasible point x^* to be a local minimizer of Problem (NLP) is that there exists multipliers

$$y_k \geq 0, \qquad k \in I^*, \qquad \mu_j \in R, \qquad j \in J, \tag{39}$$

such that

$$\nabla f^\circ(x)^* + \sum_{k\in I^*} y_k \nabla f^k(x^*) + \sum_{j\in J} \mu_j \nabla h^j(x^*) = 0, \tag{40}$$

$$d'\left[\nabla^2 f^\circ(x^*) + \sum_{k\in I^*} y_k \nabla^2 f^k(x^*) + \sum_{j\in J} \mu_j \nabla^2 h^j(x^*)\right] d \geq 0, \tag{41}$$

for every d satisfying

$$\begin{aligned} &\nabla f^k(x^*)d \leq 0, \qquad k \in I^*, \\ &\nabla h^j(x^*)d = 0, \qquad j \in J. \end{aligned} \tag{42}$$

Proof. Since x^* is a local minimum, it satisfies the necessary conditions of Theorem 3.2. However, the additional assumption (SCQ) excludes the possibility that $y_0 = 0$, for otherwise (32), (33) contradict (SCQ). Thus, $y_0 > 0$, hence can be taken equal to one (by the homogeneity of the conditions), and (32), (35) reduces to (40), (41). Moreover, (40) can have only one solution (again by the linear independence assumption), hence the multipliers are fixed. Finally, multiplying (40) by d, using (19) and (34), we see that

$$\nabla f^\circ(x^*)d = 0,$$

which justifies the absence of this relation in (42) [see (18), (19)]. □

The following example illustrates a situation in which Condition (CQ) is satisfied, but no strong constraint qualifications hold.

Example 3.1. Consider the problem given in Example 2.1. It was already shown there that, at the optimal point $x^* = (0, 0, 0)$, Kuhn–Tucker type conditions cannot hold, since no *fixed multipliers* exist. However, for the multipliers found there, $y_0(d) \not\equiv 0$; indeed, at x^* Condition (CQ) is satisfied, with $d = (-1, 1, 1)$ and arbitrary $z \in R$.

Theorem 3.2 holds in fact under strong constraint qualifications which are *weaker* than Condition (SCQ); see, e.g., Ref. 12. At any rate, Theorem 3.1 is a convenient tool to derive second-order Kuhn–Tucker type results, under these qualifications, in a similar way to the derivation of the *first-order* Kuhn–Tucker conditions through the use of the Fritz John conditions. See, e.g., Ref. 10.

Remark 3.1. In most textbooks (e.g., Refs. 4, 6, 7, 9), a result weaker than Theorem 3.3 is cited. Namely, the set of critical directions d satisfying (42) is replaced by the smaller set of *tangent directions* d satisfying

$$\begin{aligned} &\nabla f^k(x^*)d = 0, \qquad k \in I^* \\ &\nabla h^j(x^*)d = 0, \qquad j \in J. \end{aligned} \tag{43}$$

The source of this weaker result can be attributed to the traditional way of treating the active inequality constraints as equality constraints. The following example shows that replacing (42) by (43) indeed weakens the necessary conditions.

Example 3.2. Consider the problem

$$\begin{aligned} &\text{min} && f^\circ = -x_1^2 - x_1 - x_2, \\ &\text{subject to} && f^2 = \exp(-x_1) + x_2 + 1 \leqslant 0, \\ & && f^2 = -x_1 + x_2^2 \leqslant 0. \end{aligned}$$

The point $x^* = (0, 0)$ is not optimal: $x(\epsilon) = (\epsilon, -\epsilon)$, for $0 \leq \epsilon \leq 1$, is a better point. Accordingly, the necessary conditions of Theorem 3.3 are not satisfied here, since they require that

$$-3d_1^2 \geq 0, \tag{44}$$

for every d_1, d_2 such that

$$d_1 + d_2 \leq 0, \qquad -d_1 \leq 0,$$

which is clearly impossible. However, if (43) is used instead of (42), then (44) has to hold only for d_1, d_2, such that

$$d_1 + d_2 = 0, \qquad -d_1 = 0,$$

a valid statement.

4. Sufficient Conditions

In this section, we derive sufficient conditions for a feasible point x^* of Problem (NLP) to be an isolated local minimizer.

First, we derive the dual sufficient condition given in Table 1.

Theorem 4.1. *Dual Sufficient Conditions.* Consider Problem (NLP) with twice continuously differentiable functions $\{f^k: \{0\} \cup I\}$, $\{h^j: j \in J\}$. A sufficient condition for a feasible point x^* to be an isolated minimizer of Problem (NLP) is that, corresponding to every nonzero critical direction d, there exist multipliers

$$y_k \geq 0, \qquad k \in I_0^*, \qquad \mu_j \in R, \qquad j \in J, \text{ not all zero,}$$

such that (33) and (34) are satisfied and

$$d'\Big[\sum_{k \in I_0^*} y_k \nabla^2 f^k(x^*) + \sum_{j \in J} \mu_j \nabla^2 h^j(x^*)\Big] d > 0. \tag{45}$$

Proof. Suppose that x^* is not an isolated local minimizer. Then, there exists a sequence of feasible points $\{x_n\}$, $x_n \to x^*$, such that

$$f^\circ(x_n) \leq f^\circ(x^*).$$

Specifically, $\{x_n\}$ can be chosen in the form

$$x_n = x^* + t_n d_n,$$

where $\{t_n\}$ is a sequence of positive scalars converging to zero and $\{d_n\}$ is a sequence of normalized direction vectors,

$$\|d_n\| = 1.$$

For every $k \in I_0^*$, we have

$$f^k(x_n) \leq f^k(x^*);$$

hence, by Taylor expansion, for some $0 \leq \xi_n^k \leq 1$,

$$0 \geq f^k(x^* + t_n d_n) - f^k(x^*) = t_n \nabla f^k(x^*) d_n + \tfrac{1}{2} t_n^2 d_n' \nabla^2 f^k(x^* + \xi_n^k t_n d_n) d_n,$$

so that

$$0 \geq \nabla f^k(x^*) d_n + \tfrac{1}{2} t_n d_n' \nabla^2 f^k(x^* + \xi_n^k t_n d_n) d_n, \quad k \in I_0^*. \tag{46}$$

Similarly, for the equality constraints,

$$0 = \nabla h^j(x^*) d_n + \tfrac{1}{2} t_n d_n' \nabla^2 h^j(x^* + \eta_n^j t_n d_n) d_n, \qquad j \in J, \tag{47}$$

for some $0 \leq \eta_n^j \leq 1$, $j \in J$. Let $\{t_n, d_n\}_K$ be a convergent subsequence of (t_n, d_n) whose limit is thus of the form $(0, \bar{d})$, where

$$\|\bar{d}\| = 1.$$

It follows then, from (46) and (47), that

$$\nabla f^k(x^*) \bar{d} \leq 0, \quad k \in I_0^*, \qquad \nabla h^j(x^*) \bar{d} = 0, \quad j \in J;$$

i.e., $\bar{d}$ satisfies (18) and (19). Therefore, by the assumptions of the theorem, there exists multipliers $\{\bar{y}_k\}$, $\{\bar{\mu}_j\}$ satisfying, together with $\bar{d}$, (32), (33), (45). Next, multiplying (46) by $\bar{y}_k$, (47) by $\bar{\mu}_j$, and summing for all $k \in I_0^*$ and $j \in J$, one obtains

$$0 \geq \tfrac{1}{2} t_n d_n' \Big[\sum_{k \in I_0^*} \bar{y}_k \nabla^2 f^k(x^* + \xi_n^k t_n d_n) + \sum_{j \in J} \bar{\mu}_j \nabla^2 h^j(x^* + \eta_n^j t_n d_n) \Big] d_n. \tag{48}$$

Dividing by $\frac{1}{2} t_n$ and letting $\{t_n, d_n\}_K$ approach its limit, we get from (48) a contradiction to the fact that $\{\bar{y}_k\}$, $\{\bar{\mu}_j\}$, $\bar{d}$ satisfy (45). □

Remark 4.1. If one restricts the multipliers $\{y_k\}$, $\{\mu_j\}$ to be fixed (i.e., to be the same for every critical direction), then Theorem 4.1 reduces to a result due to Penissi (Ref. 5). If one further restricts y_0 to be positive, then Theorem 4.1 reduces to the widely used sufficient conditions of Fiacco and McCormick (Ref. 4, Theorem 4). Our proof of Theorem 4.1 is quite similar to the proof of Theorem 4 in Ref. 4; nevertheless, Theorem 4.1 is a strictly better sufficient condition. This is illustrated by the following example.

Example 4.1. Consider again the problem given in Example 2.1. It was already shown there that no fixed multipliers satisfy even the weak inequality (45). Thus, Penissi's result, as well as the result of Fiacco and McCormick, fail to extablish the optimality of $x^* = (0, 0, 0)$. In contrast, the sufficient conditions in Theorem 4.1 are satisfied by the multipliers

$$(y_0, y_1, y_2) = \begin{cases} (1, 1, 0), & \text{if } d_1 = 0, \\ (1, 0, 1), & \text{if } d_2 = 0, \\ (0, 1, 1), & \text{if } d_3 = 0, \\ (1, 0, 0), & \text{if } d_1 d_2 > 0, \\ (0, 1, 0), & \text{if } d_1 d_3 > 0, \\ (0, 0, 1), & \text{if } d_2 d_3 > 0. \end{cases}$$

Theorem 4.1 can be stated in an equivalent primal form given below. We note that primal forms of the above-mentioned classical sufficient conditions do not exist.

Theorem 4.2. *Primal Sufficient Conditions.* Under the assumption of Theorem 4.1, a feasible point x^* of Problem (NLP) is an isolated local minimizer if, for every critical direction d, there is no $z \in R^n$, $(z, d) \neq 0$, such that

$$\nabla f^k(x^*)z + d'\nabla^2 f^k(x^*)d \leq 0, \qquad k \in \Lambda^*(d), \tag{49}$$

$$\nabla h^j(x^*)z + d'\nabla^2 h^j(x^*)d = 0, \qquad j \in J, \tag{50}$$

where as before

$$\Lambda^*(d) \triangleq \{k: \nabla f^k(x^*)d = 0\}.$$

Proof. If $d = 0$, then the inconsistency of the system (49), (50) is a trivial sufficient condition. Thus, assume that $d \neq 0$. Consider the matrices A and B, given by their rows

$$A = \{\nabla f^k(x^*): k \in \Lambda^*(d)\}, \qquad B = \{\nabla h^j(x^*): j \in J\},$$

and consider the vectors a and b given by their components.

$$a = \{-d'\nabla^2 f^k(x^*)d: k \in \Lambda^*(d)\}, \qquad b = \{-d'\nabla^2 h^j(x^*)d: j \in J\}.$$

With these notations, the conditions of the theorem can be restated as follows: for every critical d, the system

$$Az \leq a, \qquad Bz = b$$

has no solution. By a well-known alternative theorem (e.g., Ref. 13, Theorem 1) the latter is equivalent to the *consistency* of the system

$$yA + \mu B = 0,$$

$$a'y + b'\mu > 0,$$

$$y \geqslant 0, \qquad (y, \mu) \neq 0.$$

Therefore, there exists $\{\bar{y}_k: k \in \Lambda^*(d)\}$; $\{\bar{\mu}_j: j \in J\}$ satisfying

$$\bar{y}_k \geqslant 0, \qquad k \in \Lambda^*(d), \qquad (\bar{y}, \bar{\mu}) \neq 0,$$

$$\sum_{k \in \Lambda^*(d)} \bar{y}_k \nabla f^k(x^*) + \sum_{j \in J} \mu_j \nabla h^j(x^*) = 0,$$

$$\sum_{k \in \Lambda^*(d)} \bar{y}_k(-d'\nabla^2 f^k(x^*)d) + \sum_{j \in J} \bar{\mu}_j(-d'\nabla^2 h^j(x^*)d) < 0.$$

Define

$$\bar{y}_k = 0, \qquad \text{for } k \in I_0^*, k \notin \Lambda^*(d).$$

Then

$$\{\bar{y}_k: k \in I_0^*\}, \qquad \{\bar{\mu}_j, j \in J\}$$

satisfy the sufficient conditions of Theorem 4.1. □

References

1. MANGASARIAN, O. L., and FROMOVITZ, S. *The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints*, Journal of Mathematical Analysis and Applications, Vol. 17, pp. 73–74, 1967.
2. JOHN, F., *Extremum Problem with Inequalities as Subsidiary Conditions*, Studies and Essays: Courant Anniversary Volume, Edited by K. O. Friedrichs, O. E. Neugebauer, and J. J. Stoker, John Wiley and Sons, New York, New York, 1948.
3. MCSHANE, E. J., *Sufficient Conditions for a Weak Relative Minimum in the Problem of Bolza*, Transactions of American Mathematical Society, Vol. 52, pp. 344–379, 1942.
4. FIACCO, A. V., and MCCORMICK, G. P., *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley and Sons, New York, New York, 1968.
5. PENISSI, E. J., *An Indirect Sufficiency Proof for the Problem of Lagrange with Differential Inequalities as Added Side Conditions*, Transactions of American Mathematical Society, Vol. 74, pp. 177–198, 1953.
6. MCCORMICK, G. P., *Second-Order Conditions for Constrained Minima*, SIAM Journal on Applied Mathematics, Vol. 15, pp. 641–652, 1967.

7. LUENBERGER, D. E., *Introduction to Linear and Nonlinear Programming*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1973.
8. AVRIEL, M., *Nonlinear Programming, Analysis and Methods*, Prentice Hall, Englewood Cliffs, New Jersey, 1976.
9. COLLATZ, L., and WETTERLING, W., *Optimization Problems*, Springer-Verlag, New York, New York, 1975.
10. MANGASARIAN, O. L., *Nonlinear Programming*, McGraw-Hill Book Company, New York, New York, 1969.
11. BEN-TAL, A., BEN-ISRAEL, A., and ZLOBEC, S., *Characterization of Optimality in Convex Programming without a Constraint Qualification*, Journal of Optimization Theory and Applications, Vol. 20, pp. 417–437, 1976.
12. HESTENES, M. R., *Optimization Theory, The Finite Dimensional Case*, John Wiley and Sons, New York, New York, 1975.
13. FAN, K. *On System of Linear Inequalities*, Linear Inequalities and Related Systems, Edited by H. W. Kuhn and A. W. Tucker, Princeton University Press, Princeton, New Jersey, 1956.