SIAM J. Appl. Math.
Vol. 15, No. 3, May, 1967
Printed in U.S.A.

SECOND ORDER CONDITIONS FOR CONSTRAINED MINIMA*

GARTH P. McCORMICK†

Abstract. This paper establishes two sets of "second order" conditions—one which is necessary, the other which is sufficient—in order that a vector x^* be a local minimum to the constrained optimization problem: minimize $f(x)$ subject to the constraints $g_i(x) \geqq 0$, $i = 1, \cdots, m$, and $h_j(x) = 0$, $j = 1, \cdots, p$, where the problem functions are *twice continuously differentiable*. The necessary conditions extend the well-known results, obtained with Lagrange multipliers, which apply to equality constrained optimization problems, and the Kuhn-Tucker conditions, which apply to mixed inequality and equality problems when the problem functions are required only to have continuous first derivatives. The sufficient conditions extend similar conditions which have been developed only for equality constrained problems. Examples of the applications of these sets of conditions are given.

1. Introduction. Efforts to establish conditions which determine whether or not a point solves an optimization problem have been in progress since the classical work involving Lagrange multipliers. The Lagrange multiplier approach is applied to optimization problems with equality constraints of the form given in Problem L.

Problem L. Find a vector $x^* = (x_1^*, \cdots, x_n^*)^T$ that minimizes $f(x)$ subject to

$$h_j(x) = 0, \qquad j = 1, \cdots, p.$$

Although one is interested in the true or global solution to optimization problems, in general one can only prove theorems about local solutions. A local solution is a point x^* such that in a neighborhood about that point all other points either do not satisfy the constraints of the problem or give values of the objective function greater than or equal to $f(x^*)$.

Using this definition, the basic result of Lagrange multipliers is stated in Theorem 1. (Throughout this paper, the symbol ∇ is the differentiation operator with respect to the vector x, i.e., $\nabla f(x) = (\partial f(x)/\partial x_1, \cdots, \partial f(x)/\partial x_n)^T$; the symbol ∇^2 represents the operator ∇ applied twice, $\nabla^2 f(x)$ is the $n \times n$ matrix whose i, jth element is $\partial^2 f(x)/\partial x_j \partial x_i$. For shorthand, f^* will indicate $f(x^*)$, ∇f^* will indicate $\nabla f(x^*)$; for a parameterized function, a' will represent the derivative with respect to the parameterizing variable, i.e., $f'(\theta) = df[x(\theta)]/d\theta$.) The term "differentiable" will always mean "continuously differentiable."

Theorem 1 (Lagrange multipliers) [3, p. 100]. *If the functions*

* Received by the editors April 29, 1966.

† Advanced Research Department, Research Analysis Corporation, McLean, Virginia 22101.

G. Giorgi and T.H. Kjeldsen (eds.), *Traces and Emergence of Nonlinear Programming*, 259
DOI 10.1007/978-3-0348-0439-4_12, © Springer Basel 2014

f, h_1, $\cdots$, h_p are differentiable and if x^* is a point where the vectors ∇h_1^*, $\cdots$, ∇h_p^* are linearly independent, then a necessary condition that x^* be a local minimum to Problem L *is that there exist scalars (called Lagrange multipliers) w_j^*, $j = 1, \cdots, p$, such that*

$$\nabla f^* + \sum_{j=1}^{p} w_j^* \nabla h_j^* = 0. \tag{1}$$

The proof of this theorem will be a by-product of later developments in this paper.

That the necessary conditions expressed by (1) are not always able to distinguish a local minimum from other points is seen in the following example.

Example 1. Minimize $-x_1 - x_2$ subject to

$$x_1^2 + x_2^2 - 2 = 0.$$

Using the necessary condition expressed in (1) we need examine only points $(x_1, x_2)^T$ on the circle $x_1^2 + x_2^2 - 2 = 0$ for which a scalar w_1 exists satisfying the equation

$$(-1, -1)^T + w_1(2x_1, 2x_2)^T = (0, 0)^T. \tag{2}$$

Clearly x_1 must equal x_2, and this leaves the two points $(-1, -1)^T$ and $(1, 1)^T$ as possible local minima. For the former point $w_1 = -\frac{1}{2}$ and for the latter, $w_1 = \frac{1}{2}$ would satisfy (2). Just using (1) then, there is no way to distinguish algebraically between those two points, although geometrically it is clear that $(-1, -1)^T$ is not a local minimum.

The equality constrained optimization problem, Problem L, is a special case of the constrained optimization problem, Problem I.

PROBLEM I. Find a vector x^* that minimizes $f(x)$ subject to

$$g_i(x) \geqq 0, \qquad i = 1, \cdots, m.$$

With respect to Problem I, Fritz John [1] proved a general theorem of which the following is a special case.

THEOREM 2 [1, Theorem 1, p. 188]. *If the functions f, g_1, $\cdots$, g_m are differentiable, then a necessary condition that x^* be a local minimum to Problem* I *is that there exist scalars u_0^*, u_0^*, $\cdots$, u_m^* (not all zero) such that the following inequalities and equalities are satisfied by $(x_1^*, \cdots, x_n^*, u_0^*, u_1^*, \cdots, u_m^*)$:*

$$g_i(x) \geqq 0, \qquad i = 1, \cdots, m, \tag{3}$$

$$u_0 \nabla f(x) - \sum_{i=1}^{m} u_i \nabla g_i(x) = 0, \tag{4}$$

$$(5) \qquad u_i g_i(x) = 0, \qquad i = 1, \cdots, m,$$

$$(6) \qquad u_i \geqq 0, \qquad i = 0, 1, \cdots, m.$$

The proof is omitted here.

With respect to the inequality constraints then, the variables u_i^*, $i = 0, 1, \cdots, m$, are always nonnegative.

In a later paper, Kuhn and Tucker [4] showed that if a condition, called the "first order constraint qualification," holds at x^*, then u_0^* can be taken equal to 1. The statement of the first order constraint qualification and the proof of the Kuhn-Tucker theorem are given in §2.

The failure of conditions (3) through (6) to answer questions which have proper answers is illustrated in the following example.

Example 2. Find the values of the parameter $k > 0$ for which $(0, 0)$ is a local minimum to the problem: minimize $(x_1 - 1)^2 + x_2^2$ subject to

$$-x_1 + x_2^2/k \geqq 0.$$

Using (3) through (6) the following equation must be satisfied:

$$u_0^*(-2, 0)^T - u_1^*(-1, 0)^T = (0, 0)^T.$$

Since $u_0^* = 0$ implies $u_1^* = 0$, and since Theorem 2 says they both cannot be equal to zero, it follows that u_0^* can be taken equal to 1, and $u_1^* = 2$. These values of u_0^* and u_1^* make the necessary conditions (3) through (6) valid for all values of k. But for $k = \frac{1}{4}$, $(0, 0)^T$ is not a local minimum and for $k = 3$, it is.

These two examples provide the motivation for the remainder of the paper. The theorems henceforth will be addressed to the constrained optimization problem, Problem M, containing a mixture of inequality and equality constraints.

PROBLEM M. Minimize $f(x)$ subject to

$$(7) \qquad \begin{aligned} g_i(x) &\geqq 0, \qquad i = 1, \cdots, m, \\ h_j(x) &= 0, \qquad j = 1, \cdots, p. \end{aligned}$$

In §2, using the first order constraint qualification, the Kuhn-Tucker theorem is proved. In §3, by addition of a condition called the "second order constraint qualification", additional necessary conditions are placed on a local minimum to Problem M when the problem functions are assumed twice differentiable. This is a new result although special cases have appeared elsewhere. In [5] similar results are obtained for the problem of maximizing a quadratic indefinite form subject to *linear* constraints. Next, we prove constructively that the first and second order constraint qualifications are satisfied when a regularity condition is placed on x^*. In §4, a sufficiency theorem for a point x^* to be a local minimum to Problem M is given, extending classical results which are valid only for the equality constrained problem Problem L. In §5 it is shown how the "second order" necessary and sufficient conditions solve the two examples for which the first order condition failed.

2. First order necessary conditions. Use will be made of the following lemma which is stated here without proof [2].

FARKAS' LEMMA.[1] *If every vector z (n components) which satisfies the inequality relations*

$$z^T a_i \geqq 0, \qquad i = 1, \cdots, q,$$

and the equality relations

$$z^T b_j = 0, \qquad j = 1, \cdots, r,$$

also satisfies the inequality

$$z^T c \geqq 0,$$

then there exist nonnegative scalars $t_1, \cdots, t_q$ and scalars $s_1, \cdots, s_r$ (unrestricted in sign) such that

$$c - \sum_{i=1}^{q} t_i a_i + \sum_{j=1}^{r} s_j b_j = 0.$$

We now state a condition, first introduced by Kuhn and Tucker [4, p. 483], which will be required to hold at any candidate for a local minimum.

First order constraint qualification. Let x^0 be any point satisfying the constraints of Problem M, and assume that the functions $g_1, \cdots, g_m, h_1, \cdots, h_p$ are differentiable at x^0. Then the first order constraint qualification holds at x^0, if for any nonzero vector y such that $y^T \nabla g_i(x^0) \geqq 0$ for all $i \in B^0 = \{i \mid g_i(x^0) = 0\}$, and $y^T \nabla h_j(x^0) = 0$, $j = 1, \cdots, p$, y is tangent to an arc $\alpha(\theta)$ differentiable at x^0 which is contained in the constraint region.

An arc $\alpha(\theta)$ is considered here to be a parameterized curve, differentiable when $0 \leqq \theta \leqq \epsilon$. The first order constraint qualification means that at $\alpha(0) = x^0$, $\alpha'(0) = y$.

THEOREM 3 (Kuhn-Tucker). *If the functions $f, g_1, \cdots, g_m, h_1, \cdots, h_p$ are differentiable at a point x^* and if the first order constraint qualification holds at x^*, the necessary conditions that x^* be a local minimum to the constrained optimization problem* M *are that there exist scalars $u_1^*, \cdots, u_m^*$, $w_1^*, \cdots, w_p^*$ such that (x^*, u^*, w^*) satisfies*

[1] For convenience Farkas' lemma has been restated here in a form different but equivalent to its usual form.

(8) $$g_i(x) \geqq 0, \qquad i = 1, \cdots, m,$$

(9) $$h_j(x) = 0, \qquad j = 1, \cdots, p,$$

(10) $$u_i g_i(x) = 0, \qquad i = 1, \cdots, m,$$

(11) $$u_i \geqq 0, \qquad i = 1, \cdots, m,$$

(12) $$\nabla f(x) - \sum_{i=1}^{m} u_i \nabla g_i(x) + \sum_{j=1}^{p} w_j \nabla h_j(x) = 0.$$

Proof. Let $B^* \equiv \{i \mid g_i(x^*) = 0\}$. Consider any nonzero vector y such that $y^T \nabla g_i^* \geqq 0$ for all $i \in B^*$ and $y^T \nabla h_j^* = 0$, $j = 1, \cdots, p$. By the first order constraint qualification y is the tangent of a differentiable arc, emanating from x^*, which is contained in the constraint region.

Let $\alpha(\theta)$ be that arc. Using the chain rule, the rate of change of f along $\alpha(\theta)$ at $\alpha(0) = x^*$ is

$$f'[\alpha(0)] = \alpha'(0)^T \nabla f(x^*) = y^T \nabla f(x^*).$$

By assumption, x^* is a local minimum and f must increase or remain the same along $\alpha(\theta)$. Thus, $y^T \nabla f(x^*) \geqq 0$.

The hypotheses of Farkas' lemma are satisfied by the vectors $\nabla g_i(x^*)$ for all $i \in B^*$, the vectors $\nabla h_j(x^*)$, $j = 1 \cdots p$, and the vector $\nabla f(x^*)$. Then there exist values $u_i^* \geqq 0$ for all $i \in B^*$, and w_j^*, $j = 1, \cdots, p$, such that

$$\nabla f(x^*) - \sum_{i \in B^*} u_i^* \nabla g_i(x^*) + \sum_{j=1}^{p} w_j^* \nabla h_j(x^*) = 0.$$

Let $u_i^* = 0$ for all $i \notin B^*$, and the theorem is proved.

3. Second order necessary conditions. The following may be a requirement on some point in the constraint region.

Second order constraint qualification. Let x^0 be some point satisfying the constraints of M, and assume that the functions $g_1, \cdots, g_m, h_1, \cdots, h_p$ are twice differentiable at x^0. The second order constraint qualification holds at x^0 if the following is true. Let y be any nonzero vector such that $y^T \nabla g_i^0 = 0$ for all $i \in B^0 = \{i \mid g_i(x^0) = 0\}$, and such that $y^T \nabla h_j^0 = 0$, $j = 1, \cdots, p$. Then y is the tangent of a twice differentiable arc $\alpha(\theta)$ (where $\theta \geqq 0$) along which $g_i[\alpha(\theta)] = 0$ for all $i \in B^0$, $h_j[\alpha(\theta)] = 0$ for small θ, $j = 1, \cdots, p$, and $\alpha(0) = x^0$, i.e., $\alpha'(0) = y$.

THEOREM 4 (Second order necessary conditions). *If the functions $f, g_1, \cdots, g_m, h_1, \cdots, h_p$ are twice differentiable at a point x^*, and if the first and second order constraint qualifications hold at x^*, then necessary conditions that x^* be a local minimum to the constrained optimization problem M are that there exist vectors $u^* = (u_1^*, \cdots, u_m^*)^T$ and*

$w^* = (w_1^*, \cdots, w_p^*)^T$ *such that* (8)–(12) *hold, and such that for every vector* y, *where* $y^T \nabla g_i^* = 0$ *for all* $i \in B^* = \{i \mid g_i(x^*) = 0\}$ *and such that* $y^T \nabla h_j^* = 0, j = 1, \cdots, p$, *it follows that*

$$(13) \qquad y^T\left[\nabla^2 f^* - \sum_{i=1}^{m} u_i^* \nabla^2 g_i^* + \sum_{j=1}^{p} w_j^* \nabla^2 h_j^*\right] y \geqq 0.$$

Proof. (i) The first part of the theorem is a repetition of Theorem 3 and follows because the first order constraint qualification is assumed to hold.

(ii) Let y be any nonzero vector such that

$$(14) \qquad y^T \nabla g_i^* = 0 \quad \text{for all} \quad i \in B^*,$$

and such that

$$(15) \qquad y^T \nabla h_j^* = 0, \qquad j = 1, \cdots, p.$$

(If there are none the theorem is proved.) Let $\alpha(\theta)$ be the twice differentiable arc guaranteed by the second order constraint qualification where $\alpha(0) = x^*$, $\alpha'(0) = y$. Denote $\alpha''(0)$ by z. Then

$$(16) \quad g_i''(0) = y^T(\nabla^2 g_i^*)y + z^T \nabla g_i^* = 0 \quad \text{for all} \quad i \in B^*,$$

$$(17) \quad h_j''(0) = y^T(\nabla^2 h_j^*)\, y + z^T \nabla h_j^* = 0, \qquad j = 1, \cdots, p.$$

Otherwise some g_i, $i \in B^*$, or h_j, $j = 1, \cdots, p$, would not be equal to zero along $\alpha(\theta)$. Using the u^* and w^* given by (i), (14) and (15),

$$f'(0) = y^T \nabla f^* = y^T \left(\sum_{i=1}^{m} u_i^* \nabla g_i^* - \sum_{j=1}^{p} w_j^* \nabla h_j^* \right) = 0.$$

Since x^* is a local minimum, and $f'(0) = 0$, $f''(0)$ must be $\geqq 0$. That is,

$$(18) \qquad f''(0) = y^T \nabla^2 f^* y + z^T \nabla f^* \geqq 0.$$

From this, (12), (16) and (17), it follows that

$$y^T \left[\nabla^2 f^* - \sum_{i=1}^{m} u_i^* \nabla^2 g_i^* + \sum_{j=1}^{p} w_j^* \nabla^2 h_j^* \right] y \geqq 0.$$

The following example illustrates that the first order constraint qualification can be satisfied while the second order constraint qualification fails to hold.

Example 3. Minimize x_2 subject to

$$g_1 = -x_1^9 + x_2^3 \geqq 0,$$

$$g_2 = x_1^9 + x_2^3 \geqq 0,$$

$$g_3 = x_1^2 + (x_2+1)^2 - 1 \geqq 0.$$

The solution is $\begin{pmatrix}0\\0\end{pmatrix}$. Now, $\nabla g_1^* = \nabla g_2^* = \begin{pmatrix}0\\0\end{pmatrix}$, $\nabla g_3^* = \begin{pmatrix}0\\2\end{pmatrix}$. Since all the constraints are equal to zero at $(0, 0)^T$, $B^* = (1, 2, 3)$. Any vector y such that $y \nabla g_i^* \geqq 0$ for all $i \in B^*$ must be of the form $(y_1, y_2^2)^T$. Clearly, any such vector is tangent to an arc pointing into the constraint region. Thus, the first order constraint qualification is satisfied.

Any vector y to be considered for the second order constraint qualification is of the form $\begin{pmatrix}y_1\\0\end{pmatrix}$ (where $y_1 \neq 0$ since y must be "nonzero"). Since there is no arc along which g_1, g_2 and g_3 remain equal to zero, the second order constraint fails to hold.

Note that the first order Lagrange conditions are only satisfied by $(u_1^*, u_2^*, \frac{1}{2})$, where u_1^*, u_2^* are any scalars $\geqq 0$. However,

$$\nabla^2 L(x^*, u^*) = \begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}$$

is negative definite and the second order necessary conditions do not hold. That the second order constraint qualification does not imply the first can be seen in the following example.

$$g_1 = -x_1^2 - (x_2 - 1)^2 + 1 \geqq 0,$$

$$g_2 = -x_1^2 - (x_2 + 1)^2 + 1 \geqq 0,$$

$$g_3 = x_1 \geqq 0.$$

The point $\begin{pmatrix}0\\0\end{pmatrix}$ is the solution to any problem with these three constraints. Their gradients are $\begin{pmatrix}0\\2\end{pmatrix}$, $\begin{pmatrix}0\\-2\end{pmatrix}$, $\begin{pmatrix}1\\0\end{pmatrix}$.

The second order constraint qualification is satisfied because there are no vectors orthogonal to all three gradients. The first order qualification is not since there are no arcs pointing into the region of feasibility (which is a single point). There are vectors y giving nonnegative inner products, for example, $y = \begin{pmatrix}2\\0\end{pmatrix}$.

In order to use these necessary conditions as criteria for determining if a point can be a local minimum, one must determine if the constraint qualifications are satisfied. One situation which often occurs where this can be done is given in the next theorem.

THEOREM 5 (Condition implying constraint qualifications). *Suppose the functions $g_1, \cdots, g_m, h_1, \cdots, h_p$ are twice differentiable. A sufficient condition that the first order and second order constraint qualifications be satisfied* *at a point x^* is that the vectors ∇g_i^* for all $i \in B^*$, (∇h_j^*, $j = 1, \cdots, p$) be linearly independent.*

Proof. We shall prove that the second order constraint qualification holds by constructing an arc which satisfies the hypotheses. The proof that the first order constraint qualification holds is omitted but is analogous to the one given.

Let y be any nonzero vector satisfying (14) and (15). (If none exist, the second order constraint qualification is trivially satisfied.) Let z be some vector such that

$$y^T(\nabla^2 g_i^*)\, y + z^T \nabla g_i^* = 0 \quad \text{for all} \quad i \in B^*, \tag{19}$$

$$y^T(\nabla^2 h_j^*)\, y + z^T \nabla h_j^* = 0, \qquad j = 1, \cdots, p. \tag{20}$$

Such a z exists because of the independence of the gradients. Assume there are q indices in B^* and that the inequality constraints are reordered so that $g_1, \cdots, g_q$ are those constraints. Let $c = p + q$, which, by the assumption of linear independence and the existence of a nonzero y satisfying (14) and (15), must be less than n. Let

$$M_c(\theta) = \begin{bmatrix} \dfrac{\partial g_1}{\partial x_1} & \cdots & \dfrac{\partial g_1}{\partial x_c} \\ \vdots & & \vdots \\ \dfrac{\partial g_q}{\partial x_1} & \cdots & \dfrac{\partial g_q}{\partial x_c} \\ \dfrac{\partial h_1}{\partial x_1} & \cdots & \dfrac{\partial h_1}{\partial x_c} \\ \vdots & & \vdots \\ \dfrac{\partial h_p}{\partial x_1} & \cdots & \dfrac{\partial h_p}{\partial x_c} \end{bmatrix},$$

let

$$M_{cn}(\theta) = \begin{bmatrix} \dfrac{\partial g_1}{\partial x_{c+1}} & \cdots & \dfrac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial g_q}{\partial x_{c+1}} & \cdots & \dfrac{\partial g_q}{\partial x_n} \\ \dfrac{\partial h_1}{\partial x_{c+1}} & \cdots & \dfrac{\partial h_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial h_p}{\partial x_{c+1}} & \cdots & \dfrac{\partial h_p}{\partial x_n} \end{bmatrix},$$

where each partial derivative in $M_c(\theta)$ and $M_{cn}(\theta)$ is evaluated at $\alpha(\theta)$.

We can assume without loss of generality that the vectors in $M_c(\theta)$ at $\theta = 0$ are linearly independent. Let

$$d(\theta) = \begin{bmatrix} \alpha'(\theta)^T \nabla^2 g_1 \, \alpha'(\theta) \\ \vdots \\ \alpha'(\theta)^T \nabla^2 g_q \, \alpha'(\theta) \\ \alpha'(\theta)^T \nabla^2 h_1 \, \alpha'(\theta) \\ \vdots \\ \alpha'(\theta)^T \nabla^2 h_p \, \alpha'(\theta) \end{bmatrix},$$

$$z_{cn} = \begin{bmatrix} z_{c+1} \\ \vdots \\ z_n \end{bmatrix},$$

where each matrix of second partial derivatives in $d(\theta)$ is evaluated at $\alpha(\theta)$. The arc is constructed as follows: let $\alpha(0) = x^*$, $\alpha'(0) = y$, and

$$\alpha''(\theta) = \begin{bmatrix} [M_c(\theta)]^{-1}[-M_{cn}(\theta)z_{cn} - d(\theta)] \\ z_{cn} \end{bmatrix}. \tag{21}$$

The existence of a twice differentiable function (arc) satisfying these conditions is guaranteed since the right-hand side of (21) is defined and continuous in a neighborhood about x^*. (See [6, Theorem 1.2].) It is a nontrivial arc since y was assumed not equal to zero.

That $\alpha''(0) = z$ follows by solving (19) and (20) for $(z_1, \cdots, z_c)^T$ in terms of z_{cn}. This agrees with (21) at $\theta = 0$. Let $e = (g_1, \cdots, g_q, h_1, \cdots, h_p)^T$, and let e_k be any component of e. Then

$$e_k(\theta) = e_k(0) + \theta e_k'(0) + (\theta^2/2)\, e_k''(\phi),$$

where $0 \leqq \phi \leqq \theta$ in a neighborhood about x^*. But $e(0) = 0 = e'(0)$, and

$$\begin{aligned} e''(\phi) &= d^k(\phi) + [M_c^k(\phi) : M_{cn}^k(\phi)] \begin{bmatrix} \{M_c(\phi)\}^{-1}[-M_{cn}(\phi)z_{cn} - d(\phi)] \\ z_{cn} \end{bmatrix} \\ &= 0. \end{aligned}$$

4. Second order sufficient conditions. The following conditions constitute an attempt to add as little as possible to the necessary conditions of Theorem 4 to create ones which are sufficient that a point be a local minimum.

THEOREM 6 (Second order sufficient conditions).[2] *Sufficient conditions that a point* x^* *be an isolated local minimum to the constrained optimization problem* M, *where* $f, g_1, \cdots, g_m, h_1, \cdots, h_p$ *are twice differentiable functions, are that there exist vectors* u^*, w^* *such that*

[2] A statement of this theorem when there are no inequality constraints is contained in [3, pp. 115–116].

(22) $$g_i(x^*) \geqq 0, \qquad i = 1, \cdots, m,$$

(23) $$h_j(x^*) = 0, \qquad j = 1, \cdots, p,$$

(24) $$u_i^* g_i(x^*) = 0, \qquad i = 1, \cdots, m,$$

(25) $$u_i^* \geqq 0, \qquad i = 1, \cdots, m,$$

(26) $$\nabla f^* - \sum_{i=1}^{m} u_i^* \nabla g_i^* + \sum_{i=1}^{p} w_j^* \nabla h_j^* = 0,$$

and for every nonzero vector y where $y^T \nabla g_i^* = 0$ for all $i \in D^* = \{i \mid u_i^* > 0\}$ and $y^T \nabla h_j^* = 0, j = 1, \cdots, p$, it follows that

(27) $$y^T \left[\nabla^2 f^* - \sum_{i=1}^{m} u_i^* \nabla^2 g_i^* + \sum_{j=1}^{p} w_j^* \nabla^2 h_j^* \right] > 0.$$

Proof. Assume that x^* is not an isolated local minimum. Then there exists a sequence of points $\{y^k\}$ where $\lim_{k\to\infty} y^k = x^*$ such that (i) each y^k is feasible, and (ii) $f(y^k) \leqq f(x^*)$. We can rewrite each y^k as $x^* + \delta^k s^k$ ($\delta^k > 0$), where s^k is a unit vector. We consider any limit point of the sequence $\{\delta^k, s^k\}$. Clearly any such limit point is of the form $(0, \bar{s})$, where $\bar{s}$ is a unit vector. By (i),

$$g_i(y^k) - g_i(x^*) \geqq 0 \qquad \text{for all} \qquad i \in B^*,$$

$$h_j(y^k) - h_j(x^*) = 0, \qquad j = 1, \cdots, p,$$

and by (ii),

$$f(y^k) - f(x^*) \leqq 0.$$

Dividing each equation above by δ^k, and taking the limit as $k \to \infty$ (using that sequence converging to $\bar{s}$), we have, by the assumed differentiability properties, that

(28) $$\nabla^T g_i^* \bar{s} \geqq 0 \qquad \text{for all} \qquad i \in B^*,$$

(29) $$\nabla^T h_j^* \bar{s} = 0, \qquad j = 1, \cdots, p,$$

(30) $$\nabla^T f^* \bar{s} \leqq 0.$$

We have two cases to consider, and show a contradiction arises from each of them.

(a) For the unit vector $\bar{s}$, $\nabla^T g_i^* \bar{s} > 0$ for at least one $i \in D^*$. This assumption coupled with (28), (29), (30), (24), (25) and (26) means that

$$0 \geqq \nabla^T f^* \bar{s} = \sum_{i \in D^*} u_1^* \nabla^T g_1^* \bar{s} + \sum_{j=1}^{p} w_j^* \nabla^T h_j^* \bar{s} > 0.$$

This is an impossibility.

(b) For the unit vector $\bar{s}$,

$$\nabla^T g_1{}^* \bar{s} = 0 \quad \text{for all} \quad i \in D^*.$$

Using Taylor's expansion (defining $\mathcal{L}(x, u, w) = f(x) - \sum u_1 g_1(x) + \sum w_j h_j(x)$),

$$\mathcal{L}(y^k, u^*, w^*) = \mathcal{L}(x^*, u^*, w^*) + \delta^k (s^k)^T \nabla \mathcal{L}(x^*, u^*, w^*)$$
$$+ [(\delta^k)^2/2](s^k)^T[\nabla^2 \mathcal{L}(n^k, u^*, w^*)]s^k, \tag{31}$$

where $n^k = \lambda x^* + (1 - \lambda)\delta^k(s^k)$, $0 \leqq \lambda \leqq 1$. Using properties (i) and (ii), (24), (25), (26) to reduce (31) to an inequality, then dividing by $(\delta^k)^2/2$ yields

$$0 \geqq (s^k)^T[\nabla^2 \mathcal{L}(n^k, u^*, w^*)]s^k. \tag{32}$$

Taking the limit as $k \to \infty$ yields, because of assumption (b), a statement contradicting (27).

5. Examples. The application of these theorems to the earlier examples will now be shown. In Example 1, since there is only one equality constraint, the hypotheses of Theorem 4 (by virtue of Theorem 5) are satisfied. If $(-1, -1)^T$ is a local minimum, the matrix $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ must be positive semidefinite for all vectors y such that $(y_1, y_2)(-2, -2)^T = 0$, that is, all vectors of the form $(y_1, -y_1)$. Multiplying yields $-4y_1^2 < 0$ (when $y_1 \neq 0$). Thus $(-1, -1)^T$ cannot be a local minimum. Applying the sufficiency test of Theorem 6 assures us that $(1, 1)^T$ is a local minimum. Since it is the only local minimum, it must also be the global minimum.

In Example 2, $[\nabla^2 f - \sum u_i \nabla^2 g_i]$ at $x^* = (0, 0)^T$ is $\begin{bmatrix} 2 & 0 \\ 0 & 2 - 4/k \end{bmatrix}$. Now since $\nabla g_1{}^* = (-1, 0)^T$, we need only consider vectors y of the form $(0, y_2)^T$. The number to be investigated is, therefore, $y_2^2[2 - 4/k]$. By Theorem 6, for $k > 2$, $(0, 0)^T$ is a local minimum. By Theorem 4, for $k < 2$, $(0, 0)^T$ is not a local minimum. At $k = 2$, the necessary conditions are satisfied, but not the sufficient ones.

REFERENCES

[1] F. John, *Extremum problems with inequalities as subsidiary conditions*, Studies and Essays, Interscience, New York, 1948, pp. 187–204.

[2] D. Gale, *Convex polyhedral cones and linear inequalities*, Activity Analysis of Production and Allocation, Cowles Commission Monograph 13, T. C. Koopmans, ed., John Wiley, New York, 1951.

[3] H. Hancock, *Theory of Maxima and Minima*, Dover, New York, 1960.

[4] H. W. Kuhn and A. W. Tucker, *Nonlinear programming*, Proceedings of the

Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, California, 1951, pp. 481–492.
[5] K. RITTER, *Stationary points of quadratic maximum problems*, Z. Wahrscheinlichkeitstheorie und Verw. Gebeite, 4 (1965), pp. 149–158.
[6] E. A. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.