JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 17, 37-47 (1967)

The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints

O. L. MANGASARIAN AND S. FROMOVITZ

Shell Development Company, Emeryville, California

Submitted by Lofti Zadeh

PURPOSE AND SCOPE

Optimality criteria form the foundations of mathematical programming both theoretically and computationally. In general, these criteria can be classified as either necessary or sufficient. Of course, one would like to have the same criterion be both necessary *and* sufficient. However, this occurs only under somewhat ideal conditions which are rarely satisfied in practice. In the absence of convexity, one is never assured, in general, of the sufficiency of any such optimality criterion. We are then left with only the necessary optimality criterion to face the vast number of mathematical programming problems which are not convex.

The best-known necessary optimality criterion for a mathematical programming problem is the Kuhn-Tucker criterion [1]. However, the Fritz-John criterion [2], which predates the Kuhn-Tucker criterion by about three years, is in a sense more general. In order for the Kuhn-Tucker criterion to hold, one must impose a constraint-qualification on the constraints of the problem. On the other hand, no such qualification need be imposed on the constraints in order that the Fritz John criterion hold. Moreover, the Fritz John criterion itself can be used to derive a form of the constraint qualification for the Kuhn-Tucker criterion.

Originally, Fritz John derived his conditions for the case of inequality constraints alone. If equality constraints are present and they are merely replaced by two inequality constraints, then the Fritz John original conditions become useless because every feasible point satisfies them. The new generalization of Fritz John's conditions derived in this work treats equalities *as* equalities and does not convert them to inequalities. This makes it possible to handle equalities and inequalities together.

Another contribution of the present work is a constraint qualification for equalities and inequalities together. Previous constraint qualifications treated equalities and inequalities separately, but not together. Since many realistic problems contain equalities and inequalities together, it is useful to know when the constraint qualification is indeed satisfied.

1. Introduction

Consider the following mathematical programming problem.

$$\text{Minimize } \theta(x), \text{ subject to} \tag{1.1}$$

$$\begin{aligned} g_i(x) &\leqslant 0, \qquad i \in M = \{1, 2, \ldots, m\} \\ h_j(x) &= 0, \qquad j \in K = \{1, 2, \ldots, k\}, \end{aligned}$$

where $\theta(x)$, $g_i(x)$ and $h_j(x)$ are functions defined on the n-dimensional Euclidean space E^n and have continuous first partial derivatives on E^n. For the case when the set K is empty, Fritz John [2] established the following result.

Fritz John's Necessary Optimality Conditions: $(K = \Phi)$ *If $\bar{x}$ is a solution of* (1.1) *then there exists a vector $\bar{u} = (\bar{u}_0, \bar{u}_1, \ldots, \bar{u}_m)' \in E^{m+1}$ such that*

$$\bar{u}_0 \nabla\theta(\bar{x}) + \sum_{i=1}^{m} \bar{u}_i \nabla g_i(\bar{x}) = 0 \tag{1.2}$$

$$\sum_{i=1}^{m} \bar{u}_i g_i(\bar{x}) = 0 \tag{1.3}$$

$$\bar{u} \geqslant 0, \; \bar{u} \neq 0, \tag{1.4}$$

where $\nabla\theta(\bar{x})$ denotes the n-dimensional vector of partial derivatives evaluated at $\bar{x}$, and the prime denotes the transpose.

If the set K is not empty, then the Fritz John conditions above do not apply. If one tries to eliminate the set K by replacing each equality by two inequalities, then the Fritz John conditions become useless since every feasible point can be made to satisfy these conditions. Thus, if we replace

$$h_j(x) = 0 \qquad j \in K$$

by

$$\begin{aligned} g_{j+m}(x) &\leqslant 0 \qquad j \in K \\ -g_{j+k+m}(x) &\leqslant 0 \qquad j \in K, \end{aligned}$$

then

$$\begin{aligned} \bar{u}_0 = 0 \qquad \bar{u}_i &= 0 \qquad i \in M \\ \bar{u}_{j+m} = \bar{u}_{j+k+m} &= 1 \qquad j \in K \end{aligned}$$

satisfy the Fritz John conditions for any feasible x.

One of the aims of this work will be to develop a modified set of necessary conditions that will be meaningful when the set K is not empty. These conditions will be employed to derive a constraint qualification for equality and inequality constraints *together*. No such constraint qualification has been given before. The Kuhn-Tucker constraint qualification [1] and variants thereof [3] have been given for inequality constraints alone, while the regularity condition [4,5] for the classical Lagrange multiplier condition has been given for equality constraints alone.

Vector notation will generally be used. In general, vectors will be denoted by single lower-case Latin letters, and matrices by single upper-case Latin letters. Subscripts will be used to denote components or groups of components, superscripts will be used to distinguish vectors or matrices. A vector will be a column vector. A prime (′) will indicate the transpose of a vector or matrix. Thus, the inner product of two vectors x and y will be $x'y$. The dimensionality of some vectors and matrices will not be stated explicitly, it being clear from the context.

A crucial role will be played by Motzkin's transposition theorem [6, 7, 8] which we reproduce here for convenience.

MOTZKIN'S TRANSPOSITION THEOREM. *Let A, B, and C be real constant matrices with A being nonempty. Then either the system*

$$y'A < 0 \qquad y'B \leqslant 0 \qquad y'C = 0 \tag{1.5}$$

has a solution $\bar{y}$, or the system

$$Az_1 + Bz_2 + Cz_3 = 0, \qquad z_1 \geqslant 0, \qquad z_1 \neq 0, \qquad z_2 \geqslant 0 \tag{1.6}$$

has a solution, $\bar{z}_1$, $\bar{z}_2$, $\bar{z}_3$, but never both.

In Section 2 we shall give the modified Fritz John necessary optimality conditions, and in Section 3 we shall derive the constraint qualification from these conditions.

2. THE FRITZ JOHN NECESSARY CONDITIONS IN THE PRESENCE OF EQUALITIES AND INEQUALITIES

It is convenient to start by establishing the following fundamental

LEMMA 1. *Let $f_i(x)$, $i \in L = \{1, 2,..., l\} \neq \Phi$, and $h_j(x)$, $j \in K = \{1, 2,..., k\}$, be functions defined on an open set D of E^n and have continuous first partial derivatives on D. Let the system*

$$\begin{aligned} f_i(x) &= 0, \qquad i \in L \\ h_j(x) &= 0, \qquad j \in K \end{aligned} \tag{2.1}$$

have a solution $\bar{x} \in D$, *and let the system*

$$\begin{aligned} f_i(x) &< 0, \qquad i \in L \\ h_j(x) &= 0, \qquad j \in K \end{aligned} \tag{2.2}$$

have no solution in D. *Then the system*

$$y'\nabla f_i(\bar{x}) < 0, \qquad i \in L \tag{2.3}$$

$$y'\nabla h_j(\bar{x}) = 0, \qquad j \in K \tag{2.4}$$

has no solution y *in* E^n, *provided that*

$$\nabla h_j(\bar{x}), \qquad j \in K, \text{ are linearly independent.} \tag{2.5}$$

The proof of the above lemma is somewhat lengthy and is relegated to the Appendix. Note that the case of K being empty is *not excluded* from the above lemma.

By using Motzkin's transposition theorem now, it is easy to derive a second lemma from Lemma 1. This will enable us to establish the Fritz John conditions immediately for the case of equality and inequality constraints.

LEMMA 2. *Let the assumptions of Lemma* 1 *hold. Then there exists vectors* $\bar{r} \in E^l$, $\bar{s} \in E^k$, *such that*

$$\sum_{i=1}^{l} \bar{r}_i \nabla f_i(\bar{x}) + \sum_{j=1}^{k} \bar{s}_j \nabla h_j(\bar{x}) = 0 \tag{2.6}$$

$$\bar{r} \geqslant 0 \tag{2.7}$$

$$\begin{bmatrix} \bar{r} \\ \bar{s} \end{bmatrix} \neq 0.[1] \tag{2.8}$$

PROOF. If $\nabla h_j(\bar{x})$, $j \in K$, are linearly *dependent*, then by setting $\bar{r} = 0$, conditions (2.6), (2.7), (2.8) follow from the linear dependence of $\nabla h_j(\bar{x})$, $j \in K$.

If $\nabla h_j(\bar{x})$, $j \in K$, are linearly *independent*, then we shall use Lemma 1 and Motzkin's transposition theorem. Let

$$\begin{aligned} A_i &= \nabla f_i(\bar{x}), \qquad i \in L \\ C_j &= \nabla h_j(\bar{x}), \qquad j \in K, \end{aligned}$$

where A_i are the columns of the matrix A and C_j are the columns of the

[1] This notation means that some but not all components of $\bar{r}$ and $\bar{s}$ can vanish.

matrix C. By Motzkin's transposition theorem, with B empty, conditions (2.6), (2.7), (2.8) follow from (1.6) by setting $\bar{r} = \bar{z}_1$ and $\bar{s} = \bar{z}_3$. Q.E.D.

Note that the case of K being empty is *not* excluded from the above lemma.

We are now ready to establish the main result of this work.

The generalized Fritz John Necessary Conditions: If $\bar{x}$ is a solution of (1.1) *then there exists vectors $\bar{u} = (\bar{u}_0, \bar{u}_1, ..., \bar{u}_m)' \in E^{m+1}$, $\bar{v} = (\bar{v}_1, \bar{v}_2, ..., \bar{v}_k)' \in E^k$ such that*

$$\bar{u}_0 \nabla\theta(\bar{x}) + \sum_{i=1}^{m} \bar{u}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{k} \bar{v}_j \nabla h_j(\bar{x}) = 0 \tag{2.9}$$

$$\sum_{i=1}^{m} \bar{u}_i g_i(\bar{x}) = 0 \tag{2.10}$$

$$\bar{u} \geqslant 0 \tag{2.11}$$

$$\begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} \neq 0 \tag{2.12}$$

PROOF. Let

$$\bar{M} = \{i \mid i \in M, g_i(\bar{x}) = 0\} \tag{2.13}$$

and

$$D = \{x \mid x \in E^n, g_i(x) < 0, i \in M - \bar{M}\} \tag{2.14}$$

Since $M - \bar{M}$ is finite, and $g_i(x)$ are continuous, D is an open set in E^n. Let

$$f_1(x) = \theta(x) - \theta(\bar{x}) \tag{2.15}$$

$$f_i(x) = g_{p_i}(x), \qquad p_i \in \bar{M}, \qquad i = 2, 3, ..., l. \tag{2.16}$$

Note that since $f_1(x) = \theta(x) - \theta(\bar{x})$, the set $L = \{1, 2, ..., l\}$ is not empty.

Now the system

$$\begin{aligned} f_i(x) &= 0, \qquad i \in L \\ h_j(x) &= 0, \qquad j \in K \end{aligned} \tag{2.17}$$

has a solution $\bar{x} \in D$, but the system

$$\begin{aligned} f_i(x) &< 0, \qquad i \in L \\ h_j(x) &= 0, \qquad j \in K \end{aligned} \tag{2.18}$$

has no solution in D. For if (2.18) did have a solution $\hat{x} \in D$, then we would have that

$$\begin{aligned} \theta(\hat{x}) - \theta(\bar{x}) &< 0 \\ g_i(\hat{x}) &< 0, \qquad i \in M \\ h_j(\hat{x}) &= 0, \qquad j \in K, \end{aligned}$$

which contradicts the assumption that $\theta(\bar{x})$ is the minimum of $\theta(x)$ on the feasible set

$$S = \{x \mid x \in E^n, g_i(x) \leqslant 0, i \in M, h_j(x) = 0, j \in K\}. \tag{2.19}$$

By Lemma 2 it follows then that there exist vectors

$$\bar{r} = (\bar{r}_0, \bar{r}_{p_2}, \bar{r}_{p_3}, \ldots, \bar{r}_{p_l})' \in E^l$$

and

$$\bar{s} = (\bar{s}_1, \bar{s}_2, \ldots, \bar{s}_k)' \in E^k$$

such that

$$\bar{r}_0 \nabla\theta(\bar{x}) + \sum_{i=2}^{l} \bar{r}_{p_i} \nabla g_{p_i}(\bar{x}) + \sum_{j=1}^{k} \bar{s}_j \nabla h_j(\bar{x}) = 0, \qquad p_i \in \bar{M} \tag{2.20}$$

$$\bar{r} \geqslant 0 \tag{2.21}$$

$$\begin{bmatrix} \bar{r} \\ \bar{s} \end{bmatrix} \neq 0 \tag{2.22}$$

By defining $\bar{u} \in E^{m+1}$ and $\bar{v} \in E^k$ as follows

$$\bar{u}_0 = \bar{r}_0$$

$$\bar{u}_i = \begin{cases} 0 & i \in M - \bar{M}, \\ \bar{r}_i & i \in \bar{M} \end{cases}$$

$$\bar{v} = \bar{s},$$

we have

$$\sum_{i=1}^{m} \bar{u}_i g_i(\bar{x}) = 0. \tag{2.23}$$

Conditions (2.20), (2.21), (2.22) may be rewritten as

$$\bar{u}_0 \nabla\theta(\bar{x}) + \sum_{i=1}^{m} \bar{u}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{k} \bar{v}_j \nabla h_j(\bar{x}) = 0 \tag{2.24}$$

$$\bar{u} \geqslant 0 \tag{2.25}$$

$$\begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} \neq 0. \quad \text{Q.E.D.} \tag{2.26}$$

An interesting question may be raised in connection with the previous result which is this: What points in the feasible set S (2.19), not necessarily the minimum points of $\theta(x)$, satisfy the generalized Fritz John conditions (2.9) to (2.12)? This question can be partially answered by the following corollary which can be easily deduced from the proof of Lemma 2 and the proof of the generalized Fritz John conditions.

COROLLARY. *Let $\bar{x}$ be any point (not necessarily the minimum point) in the feasible set S. Then if either*

$$\nabla h_j(\bar{x}), \qquad j \in K, \text{ are linearly dependent} \tag{2.27}$$

or if the system

$$\left.\begin{cases} \theta(x) - \theta(\bar{x}) < 0 & \\ g_i(x) < 0, & i \in M \\ h_j(x) = 0, & j \in K \end{cases}\right\} \text{ has no solution } x \in E^n, \tag{2.28}$$

then the generalized Frits John conditions are satisfied at $\bar{x}$.

One may use conditions (2.27) and (2.28) to impose a regularity condition of the following type: All points $\bar{x}$ in S satisfying (2.27) or (2.28) must be minimum points.

3. CONSTRAINT QUALIFICATION FOR EQUALITY AND INEQUALITY CONSTRAINTS

It is well known that necessary optimality conditions such as the Kuhn-Tucker [1] or the classical Lagrange multiplier conditions [4, 5] require some sort of a constraint qualification in order for them to be valid. Such constraint qualifications have always been given for inequality constraints alone [1, 3] or for equality constraints alone [4, 5]. By using the generalized Fritz John conditions derived in the previous section, it is possible to derive a constraint qualification for *both* equality and inequality constraints. (Cottle [9] has derived a constraint qualification for inequality constraints alone by using Fritz John's original necessary conditions.)

THE GENERALIZED CONSTRAINT QUALIFICATION. *Let $\bar{x}$ be a solution of (1.1). A sufficient condition for the Kuhn-Tucker necessary optimality conditions* [1]

$$\nabla\theta(\bar{x}) + \sum_{i=1}^{m} \bar{u}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{k} \bar{v}_j \nabla h_j(\bar{x}) = 0 \tag{3.1}$$

$$\sum_{i=1}^{m} \bar{u}_i g_i(\bar{x}) = 0 \tag{3.2}$$

$$\bar{u}_i \geqslant 0, \qquad i \in M, \tag{3.3}$$

to hold is that there exists a vector $\bar{y} \in E^n$ such that

$$\bar{y}'\nabla g_i(\bar{x}) < 0 \qquad i \in \bar{M} = \{i \mid i \in M, g_i(\bar{x}) = 0\} \neq \Phi \tag{3.4}$$

$$\bar{y}'\nabla h_j(\bar{x}) = 0, \qquad j \in K, \tag{3.5}$$

and that

$$\nabla h_j(\bar{x}), \qquad j \in K, \text{ are lineary independent.} \tag{3.6}$$

For the case when $\bar{M}$ is empty, condition (3.6) *alone is a sufficient constraint qualification.*[2]

PROOF. For the case when $\bar{M} = \Phi$, it is obvious that the Fritz John conditions (2.9) to (2.12) cannot hold with $\bar{u}_0 = 0$ if (3.6) is valid. Hence, $\bar{u}_0 > 0$ and the Kuhn-Tucker conditions (3.1) to (3.3) follow from the Fritz John conditions (2.9) to (2.12).

For the case of $\bar{M} \neq \Phi$, define matrices A and C whose columns A_i and C_j are given by

$$A_i = \nabla g_i(\bar{x}), \qquad i \in \bar{M} \tag{3.7}$$

$$C_j = \nabla h_j(\bar{x}), \qquad j \in K. \tag{3.8}$$

It follows then from (3.4) and (3.5) that (1.5) has a solution, with B empty. Hence, by Motzkin's transposition theorem, the system

$$Az_1 + Cz_3 = 0, \qquad z_1 \geqslant 0, \qquad z_1 \neq 0,$$

has no solution z_1, z_3, or

$$\sum_{i \in \bar{M}} z_{1i} \nabla g_i(\bar{x}) + \sum_{j=1}^{k} z_{2j} \nabla h_j(\bar{x}) = 0 \tag{3.9}$$

$$z_1 \geqslant 0, \qquad z_1 \neq 0, \tag{3.10}$$

has no solution. Now the Fritz John conditions (2.9) to (2.12) hold. If $\bar{u}_0 = 0$, then $(\bar{u}_1, ..., \bar{u}_m) \geqslant 0$ and $(\bar{u}_1, ..., \bar{u}_m) \neq 0$ because of (3.6) and (2.9), hence

$$\bar{z}_{1i} = \bar{u}_i, \qquad i \in \bar{M}$$

$$\bar{z}_2 = \bar{v}$$

solves (3.9), (3.10), which is a contradiction. Hence, $\bar{u}_0 > 0$ and the Kuhn-Tucker condition (3.1) to (3.3) follow from the Fritz John conditions (2.9) to (2.12). Q.E.D.

[2] Note that (3.5) and (3.6) are not necessarily contradictory for (3.5) implies the linear dependence of the *rows* of $[\nabla h_1(\bar{x}), ..., \nabla h_k(\bar{x})]$.

It should be remarked that condition (3.6) is equivalent to the regularity condition of the Lagrange multiplier method for equality constraints [4, 5].

A geometric interpretation of conditions (3.4) and (3.5) can be given as follows. The gradients of the active (that is $g_i(\bar{x}) = 0$) inequality constraints at $\bar{x}$ from a pointed[3] cone, and there exists a vector in this cone that is tangent to the surface formed by the equality constraints.

Appendix

Proof of Lemma 1. The proof will be by contradiction. We shall assume that (2.1) has a solution $\bar{x} \in D$, so

$$f_i(\bar{x}) = 0, \qquad i \in L \tag{1}$$

$$h_j(\bar{x}) = 0, \qquad j \in K; \tag{2}$$

that (2.3), (2.4) have a solution $\bar{y} \in E^n$, so

$$\bar{y}' \nabla f_i(\bar{x}) < 0, \qquad i \in L \tag{3}$$

$$\bar{y}' \nabla h_j(\bar{x}) = 0, \qquad j \in K; \tag{4}$$

and that (2.5) holds. We shall then produce an $\tilde{x} \in E^n$ such that

$$f_i(\tilde{x}) < 0, \qquad i \in L \tag{5}$$

$$h_j(\tilde{x}) = 0, \qquad j \in K, \tag{6}$$

which contradicts (2.2). (For the case of $k = n$, the proof of Lemma 1 is trivial, because (2.5) and (2.4) imply that $y = 0$, and hence (2.3) has no solution. The case of $k > n$ is excluded by (2.5). So we shall only consider the case $k < n$.)

Let $f(x)$ denote the l-by-1 vector mapping from E^n into E^l defined by $f_1(x), \ldots, f_l(x)$, and $h(x)$ the k-by-1 vector mapping from E^n into E^k defined by $h_1(x), \ldots, h_k(x)$. Let $\nabla f(x)$ be the n-by-l matrix of partial derivatives $\partial f_i(x)/\partial x_j$, $i \in L$, $j \in \{1, \ldots, n\}$. Similarly define the n-by-k matrix $\nabla h(x)$, and the other matrices of partial derivatives of vector valued functions appearing below.

From implicit function theory [5, 10], (2) and (2.5) it follows that there exist:

[3] A cone is pointed [9] if there exists a vector which makes an acute angle ($< \pi/2$) with all the vectors of the cone.

a partition (x_I, x_{II}) of x, such that $x_I \in E^{n-k}$, $x_{II} \in E^k$, a neighborhood U of $\bar{x}_I$ in E^{n-k}, and a differentiable mapping $e: U \to E^k$ such that

$$\bar{x}_{II} = e(\bar{x}_I) \tag{7}$$

$$h(x_I, e(x_I)) = 0, \qquad \text{for all} \qquad x_I \in U, \tag{8}$$

and

$$\nabla_{x_{II}} h(\bar{x}) \text{ is nonsingular.} \tag{9}$$

Let (y_I, y_{II}) be the partition of y corresponding to (x_I, x_{II}). By the differentiability of $h(x)$ and $e(x)$, (8) and the chain rule [10]

$$\nabla_{x_I} h(\bar{x}) + \nabla_{x_I} e(\bar{x}_I) \nabla_{x_{II}} h(\bar{x}) = 0 \tag{10}$$

Premultiplying (10) by $\bar{y}_I'$ we get

$$\bar{y}_I' \nabla_{x_I} h(\bar{x}) + \bar{y}_I' \nabla_{x_I} e(\bar{x}_I) \nabla_{x_{II}} h(\bar{x}) = 0. \tag{11}$$

By (4) we have that

$$\bar{y}_I' \nabla_{x_I} h(\bar{x}) + \bar{y}_{II}' \nabla_{x_{II}} h(\bar{x}) = 0. \tag{12}$$

Hence by (11), (12) and (9) we get

$$\bar{y}_{II}' = \bar{y}_I' \nabla_{x_I} e(\bar{x}_I). \tag{13}$$

By the differentiability of $e(x_I)$, there exists a $\delta_1 > 0$, such that for all $\delta < \delta_1$

$$e(\bar{x}_I + \delta \bar{y}_I) = e(\bar{x}_I) + (\delta \nabla_{x_I} e(\bar{x}_I))' \bar{y}_I + \delta c(\bar{x}_I, \delta \bar{y}_I) \| \bar{y}_I \|, \tag{14}$$

where $\| \bar{y}_I \|$ denotes the Euclidean norm $(\bar{y}_I' \bar{y}_I)^{\frac{1}{2}}$ and $c(\bar{x}_I, \delta \bar{y}_I)$ is a k-by-1 vector mapping from E^{n-k} into E^k such that $\lim_{\delta \to 0} c(\bar{x}_I, \delta \bar{y}_I) = 0$. From (13) and (14) we get that

$$e(\bar{x}_I + \delta \bar{y}_I) = e(\bar{x}_I) + \delta \bar{y}_{II} + \delta c(\bar{x}_I, \delta \bar{y}_I) \| \bar{y}_I \|, \qquad \text{for all} \qquad \delta < \delta_1. \tag{15}$$

Again by the differentiability of $f(x)$, (15) and (7) we have for $\delta < \delta_1$

$$\begin{aligned} f(\bar{x}_I + \delta \bar{y}_I, e(\bar{x}_I + \delta \bar{y}_I)) &= f(\bar{x}_I + \delta \bar{y}_I, e(\bar{x}_I) + \delta \bar{y}_{II} + \delta c(\bar{x}_I, \delta \bar{y}_I) \| \bar{y}_I \|) \\ &= f(\bar{x}_I, \bar{x}_{II}) + \delta \{ (\nabla_{x_I} f(\bar{x}_I, \bar{x}_{II}))' \bar{y}_I \\ &\quad + (\nabla_{x_{II}} f(\bar{x}_I, \bar{x}_{II}))' (\bar{y}_{II} + c(\bar{x}_I, \delta \bar{y}_I) \| \bar{y}_I \|) \\ &\quad + b(\bar{x}_I, \bar{x}_{II}; \delta \bar{y}_I, \delta \bar{y}_{II} + \delta c(\bar{x}_I, \delta \bar{y}_I) \| \bar{y}_I \|) \\ &\quad \| \bar{y}_I, \bar{y}_{II} + c(\bar{x}_I, \delta \bar{y}_I) \| \bar{y}_I \| \| \}, \end{aligned} \tag{16}$$

where b is an l-by-1 vector mapping from E^n into E^l such that $\lim_{\delta \to 0} b = 0$. Since $\lim_{\delta \to 0} c = 0$, and

$$(\nabla_{x_I} f(\bar{x}_I, \bar{x}_{II}))' \bar{y}_I + (\nabla_{x_{II}} f(\bar{x}_I, \bar{x}_{II}))' \bar{y}_{II} < 0, \qquad \text{(by (3))},$$

it follows that there exists a $\delta_2 > 0$, such that for all δ, $0 < \delta < \delta_2$, the expression in the curly brackets in (16) is strictly negative. And since $f(\bar{x}_I, \bar{x}_{II}) = f(\bar{x}) = 0$, we then have from (16) that

$$f(\bar{x}_I + \delta \bar{y}_I, e(\bar{x}_I + \delta \bar{y}_I)) < 0, \qquad \text{for all} \qquad \delta : 0 < \delta < \delta_2, \tag{17}$$

and by (8)

$$h(\bar{x}_I + \delta \bar{y}_I, e(\bar{x}_I + \delta \bar{y}_I)) = 0, \qquad \text{for all} \qquad \delta : 0 < \delta < \delta_2. \tag{18}$$

By setting $\tilde{x}_I = \bar{x}_I + \delta \bar{y}_I$, $\tilde{x}_{II} = e(\bar{x}_I + \delta \bar{y}_I)$ for some $\delta : 0 < \delta < \delta_2$, relations (17) and (18) give the desired contradiction.

The case $K \neq \Phi$ goes through in a similar manner as above but without using implicit function theory.

References

1. H. W. Kuhn and A. W. Tucker. Nonlinear programming. *In* "Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability" (J. Neyman, ed.), pp. 481-492. Univ. of Calif. Press, Berkeley, Calif., 1951.
2. F. John. Extremum problems with inequalities as side conditions. *In* "Studies and Essays, Courant Anniversary Volume" (K. O. Friedrichs, O. E. Neugebauer and J. J. Stoker, eds.), pp. 187-204. Wiley (Interscience), New York, 1948.
3. K. J. Arrow, L. Hurwicz, and H. Uzawa. Constraint qualification in maximization problems. *Naval Research Logistics Quarterly* **8** (1961), 175-191.
4. R. Courant. "Differential and Integral Calculus," Vol. II, Chapter III. Wiley (Interscience), New York, 1936.
5. T. M. Apostol. "Mathematical Analysis." pp. 146-157. Addison-Wesley, Reading, Mass., 1957.
6. T. S. Motzkin. Two consequences of the transposition theorem of linear Inequalities. *Econometrica* **19** (1951), 184-185.[4]
7. M. L. Slater. A note on Motzkin's transposition theorem. *Econometrica* **19** (1951), 185-187.
8. A. W. Tucker. Dual systems of homogeneous linear relations. *In* "Linear Inequalities and Related Systems," pp. 3-18, Corollary 2A. Annals of Mathematics Studies No. 38, Princeton University Press, Princeton, New Jersey, 1956.
9. R. W. Cottle. A theorem of Fritz John in mathematical programming. RAND Memorandum RM-3858-PR, October, 1963.
10. W. H. Fleming. "Functions of Several Variables," pp. 116-120. Addison-Wesley, Reading, Mass., 1965.

[4] The first inequality appearing in this paper should be reversed.