PROCEEDINGS *of the* THIRD BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY

Held at the Statistical Laboratory
University of California
26-31 *December, 1954*
July and August, 1955

VOLUME V

CONTRIBUTIONS TO
ECONOMETRICS, INDUSTRIAL RESEARCH, AND PSYCHOMETRY

EDITED BY JERZY NEYMAN

For further effective support of the Symposium thanks must be given the National Science Foundation, the United States Air Force Research and Development Command, the United States Army Office of Ordnance Research, and the United States Navy Office of Naval Research.

UNIVERSITY OF CALIFORNIA PRESS
BERKELEY AND LOS ANGELES
1956

REDUCTION OF CONSTRAINED MAXIMA TO SADDLE-POINT PROBLEMS

KENNETH J. ARROW AND LEONID HURWICZ
STANFORD UNIVERSITY, UNIVERSITY OF MINNESOTA

1. Introduction

1.1. The usual applications of the method of Lagrangian multipliers, used in locating constrained extrema (say maxima), involve the setting up of the *Lagrangian expression*,

$$\phi(x, y) = f(x) + y'g(x), \tag{1}$$

where $f(x)$ is being (say) maximized with respect to the (vector) variable $x = \{x_1, \cdots, x_N\}$, subject to the constraint $g(x) = 0$, where $g(x)$ maps the points of the N-dimensional x-space into an M-dimensional space, and $y = \{y_1, \cdots, y_M\}$ is the Lagrange multiplier (vector). Here, $\{\quad\}$ indicates a column vector; the prime indicates transposition, so that y' is a row vector.

The essential step of the customary procedure is the solution for x, as well as y, of the pair of (vector) equations,

$$\phi_x(x, y) = 0, \; g(x) = 0, \tag{2}$$

where $\phi_x(x, y) = \{\partial\phi(x, y)/\partial x_1, \cdots, \partial\phi(x, y)/\partial x_N\}$. Let $(\bar{x}, \bar{y})$ be the solutions of equations (2), while $\hat{x}$ maximizes $f(x)$ subject to $g(x) = 0$. Then, under suitable restrictions,

$$\bar{x} = \hat{x}. \tag{3}$$

1.2. In [1] Kuhn and Tucker treat the related problem of maximizing $f(x)$ subject to the constraints[1] $g(x) \geqq 0$, $x \geqq 0$, where, for an arbitrary K-dimensional vector $a = \{a_1, \cdots, a_K\}$, the relation $a \geqq 0$ is here defined to mean $a_k \geqq 0$ for $k = 1, \cdots, K$. Another definition of vectorial inequalities, permitting greater generality of treatment, will be used in later sections of this paper. There we shall treat directly the class of situations where $f(x)$ is to be maximized subject to $g^{(1)}(x) \geqq 0$, $g^{(2)}(x) = 0$, $x^{[1]} \geqq 0$, $x^{[2]}$ not restricted as to sign, $x = \{x^{[1]}, x^{[2]}\}$.

Denote by C_g the set of all x satisfying the constraints $g(x) \geqq 0$, $x \geqq 0$. The two results stated below are of fundamental importance for the problem considered.

(A) (See theorem 1 [1].) Let g satisfy the following condition (called Constraint

Most of the work of this paper was done under the auspices of The RAND Corporation, with additional support and assistance from the Cowles Commission for Research in Economics and the Office of Naval Research.

[1] In [1] our f and g are respectively written as g and F. The symbol in [1] for the Lagrange multiplier (our y) is u.

Qualification, here abbreviated as C.Q.).[2] If $\tilde{x}$ is a boundary point of C_g and x satisfies the relations,

$$\tilde{g}_x^{a}(x - \tilde{x}) \geqq 0\,, \tag{4}$$

$$x^b - \tilde{x}^b \geqq 0\,, \tag{5}$$

where "$\sim$" over a symbol denotes its evaluation at $x = \tilde{x}$, $g = \{g^a, g^\beta\}$, $\tilde{g}^a = 0$, $\tilde{g}^\beta > 0$, $x = \{x^a, x^b\}$, $\tilde{x}^a > 0$, and $\tilde{x}^b = 0$, then there exists a differentiable vector-valued function ψ of the real variable θ whose domain is the closed interval (0, 1) and the range is in C_g; that is, $x = \psi(\theta)$, such that $\psi(0) = \tilde{x}$ and $\psi'(0) = \lambda(x - \tilde{x})$ for some positive scalar λ.

Under this condition, if all derivatives used below exist and if $\bar{x}$ maximizes $f(x)$ for $x \in C_g$, there exists y satisfying the conditions

$$\bar{x} \geqq 0, \quad \bar{\phi}_x \leqq 0, \quad \bar{x}'\bar{\phi}_x = 0\,, \tag{6}$$

$$\bar{y} \geqq 0, \quad \bar{\phi}_y \geqq 0, \quad \bar{y}'\bar{\phi}_y = 0\,, \tag{7}$$

where $\bar{\phi}_x$ and $\bar{\phi}_y$ are partial (vector) derivatives of the Lagrangian expression (1) evaluated at $(\bar{x}, \bar{y})$.

(B). (See theorem 3 [1].) If the hypotheses specified in (A) hold and, in addition, the functions $f(x)$, $g_m(x)$, $m = 1, \cdots, M$ are *concave*,[3] there exists a pair $(\bar{x}, \bar{y})$, satisfying conditions (6) and (7), such that (x, y) is a *nonnegative saddle-point* (NNSP) of $\phi(x, y)$, that is,

$$\phi(x, \bar{y}) \leqq \phi(\bar{x}, \bar{y}) \leqq \phi(\bar{x}, y) \quad \text{for all} \quad x \geqq 0, \quad y \geqq 0\,; \tag{8}$$

furthermore, any NNSP $(\bar{x}, \bar{y})$ of $\phi(x, y)$ has the property that x maximizes $f(x)$ in C_g. According to lemma 1 [1], conditions (6), (7) are implied by (8) regardless of the nature of $\phi(x, y)$, that is, even if $\phi(x, y)$ is not given by (1).

2. A modified Lagrangian approach

2.1. Because of the interesting game theoretical and economic implications of the theorem in (B), section 1.2 (which the authors will study elsewhere), the question arises as to the possibility of similar results when some of the conditions of the theorem are relaxed.

It turns out that results of such nature can be obtained, though not without some sacrifices. The relaxation is primarily with regard to the convexity assumptions which fail to hold in some important economic applications (the case of "increasing returns"). The main sacrifices are (1) the Lagrangian expression is modified, and (2) the results are proved only locally.

The results are presented below in the form of three theorems. Theorem 1 is auxiliary in nature; theorems 2 and 3 together imply the existence of a local nonnegative saddle-

[2] This restriction "is designed to rule out singularities on the boundary of the constraint set, such as an outward-pointing 'cusp'" (see p. 483 in [1]). It should be noted, however, that because of (4), C.Q. is a property of g, not merely of C_g. Thus $g(x) \equiv -(x - 1)^3$, x one-dimensional, lacks C.Q., while $g(x) \equiv -(x - 1)$, with the same C_g, does have it.

[3] A function $f(x)$ is said to be concave if

$$(1 - \theta)f(x^0) + \theta f(x) \leqq f[(1 - \theta)x^0 + \theta x]$$

for all $0 \leqq \theta \leqq 1$ and all x^0 and x in the region where $f(x)$ is defined (see [1], pp. 10–11).

point for the modified Lagrangian expression. Theorem 3 shows this saddle-point to be of the type leading to convergence in gradient procedures described by the authors in [3].

The notation differs in some detail from that introduced in section 1. To facilitate reading, some notational principles are stated in 2.2.1; the main symbols used are listed in sections 2.2.2 and 2.3.4.

2.2.1. *Some principles of notation.* A K-dimensional column vector $\{a_1, a_2, \cdots, a_K\}$ is denoted by a; dim a denotes the number of components in a. If A is a matrix, A' is its transpose. Hence, in particular, a' is a row vector and $a'b \equiv \sum_{k=1}^{K} a_k b_k$ is the inner product of the vectors a and b; $a \cdot b$ is an alternative, and sometimes more convenient, notation for $a'b$.

$[a_1, a_2, \cdots, a_K]$ is the finite (unordered) set whose elements are $a_1, a_2, \cdots, a_K$. $A \sim B$ is the set of all elements in A but not in B (the set-theoretic difference).

$\{x \mid p_x\}$ denotes the set of all x possessing the property p_x.

If

(9) $$c(a) = \{c_1(a), c_2(a), \cdots, c_P(a)\},$$

(10) $$a = \{a_1, a_2, \cdots, a_K\},$$

then

(11) $$c_a \equiv c_a(a) = \left\| \frac{\partial c_p}{\partial a_k} \right\|, \qquad p = 1, 2, \cdots, P; \qquad k = 1, 2, \cdots, K.$$

Further, $\bar{c}$, $\bar{c}_a$ denote, respectively, $c(a)$ and $c_a(a) \equiv c_a$ evaluated at $a = \bar{a}$.

If $\psi(a, b)$ is a real-valued (scalar) function of the vectors $a = \{a_1, a_2, \cdots, a_K\}$, $b = \{b_1, b_2, \cdots, b_R\}$, then

(12) $$\psi_{ab} = \left\| \frac{\partial^2 \psi}{\partial a_k \partial b_r} \right\|, \qquad k = 1, 2, \cdots, K; \qquad r = 1, 2 \cdots, R,$$

where $\bar{\psi}_{ab}$ denotes ψ_{ab} evaluated at $(\bar{a}, \bar{b})$.

$S_\rho(x^0) = \{x \mid d(x, x^0) \leqq \rho\}$ where $d(x', x'')$ denotes the Euclidean distance between x' and x''.

2.2.2. *Some symbols used.*

(N.1.1) $$x = \{x_1, x_2, \cdots, x_N\}.$$

X is the Euclidean N-space of the x's.

$$\mathcal{N} = [1, 2, \cdots, N].$$

$\mathcal{N}'$ is a fixed (possibly empty, not necessarily proper) subset of $\mathcal{N}$. As will be seen in (N.1.4), the elements of $\mathcal{N}'$ are the indices of the components of $x^{[1]}$ as defined in the first paragraph of section 1.2.

(N.1.2) $$z = \{z_1, z_2, \cdots, z_M\}.$$

Z is the Euclidean M-space of the z's.

$$\mathcal{M} = [1, 2, \cdots, M].$$

$\mathcal{M}'$ is a fixed (possibly empty, not necessarily proper) subset of $\mathcal{M}$. As will be seen from (N.1.4), (N.2), (N.3), the elements of $\mathcal{M}'$ are the indices of the components of $g^{(1)}$ as defined in the first paragraph of section 1.2; the elements of $\mathcal{M} \sim \mathcal{M}'$ are the indices of $g^{(2)}$ (see same paragraph); g will be defined as $\{g^{(1)}, g^{(2)}\}$.

(N.1.3) $$y = \{y_1, y_2, \cdots, y_M\} .$$

Y is the Euclidean M-space of the y's. Here Y is the space of the real-valued linear functions on Z. Even in the Euclidean case it is convenient to distinguish between the two, since our definitions of nonnegativity in the two spaces differ.

(N.1.4) $$x \geqq 0 \text{ means } \begin{cases} x_n \geqq 0 \text{ for } n \in \mathcal{N}' . \\ x_n \text{ unrestricted as to sign for } n \notin \mathcal{N}' . \end{cases}$$

X^+ is the set of all $x \geqq 0$.

$$z \geqq 0 \text{ means } \begin{cases} z_m \geqq 0 \text{ for } m \in \mathcal{M}' . \\ z_m = 0 \text{ for } m \notin \mathcal{M}' . \end{cases}$$

$$y \geqq 0 \text{ means } \begin{cases} y_m \geqq 0 \text{ for } m \in \mathcal{M}' . \\ y_m \text{ unrestricted as to sign for } m \notin \mathcal{M}' . \end{cases}$$

For any vector $a = \{a_1, a_2, \cdots, a_k\}$,

$$a = 0 \text{ means } a_1 = 0, a_2 = 0, \cdots, a_K = 0 ;$$

$$a > 0 \text{ means } a_1 > 0, a_2 > 0, \cdots, a_K > 0 ;$$

$$a < 0 \text{ means } -a > 0 .$$

(N.2.1) $'g$ is a function on X^+ to Z. Hence $'g(x) = \{'g_1(x), 'g_2(x), \cdots, 'g_M(x)\}$ where the $'g_m$, $m \in \mathcal{M}$ are real-valued functions.

(N.2.2) We shall find it convenient to work with some of the $'g_m$, $m \in \mathcal{M}'$ replaced by their negatives. More precisely, we write

$$g_m = \begin{cases} 'g_m \text{ if } m \in \mathcal{M} \sim \mathcal{M}^- \\ -'g_m \text{ if } m \in \mathcal{M}^- , \end{cases}$$

where $\mathcal{M}^- \subseteq \mathcal{M} \sim \mathcal{M}'$ will be defined in section 2.3.4.

$$g = \{g_1, g_2, \cdots, g_M\} .$$

Note. Since $\mathcal{M}^- \subseteq \mathcal{M} \sim \mathcal{M}'$, it is seen that the conditions

$$'g(x) \geqq 0, \quad g(x) \geqq 0$$

are equivalent. For practical purposes, one could consider the problem as given directly in terms of g, rather than $'g$. We start with $'g$, however, in order to avoid the impression of a loss of generality in connection with the assumptions of section 2.3.4.

(N.3) $$C_g = \{x \,|'g(x) \geqq 0, \, x \geqq 0\} \equiv \{x \,|g(x) \geqq 0, \, x \geqq 0\}$$

(the "constraint set").

(N.4) f is a real-valued function on X^+ (the "maximand") .

(N.5) $$O_{fg} = \{x' \,|x' \in C_g \text{ and } f(x) \leqq f(x') \text{ for all } x \in C_g\}$$

(the "optimal set").

(N.6) $x = \{x^{(1)}, x^{(2)}\}$ where

$\mathcal{N}^{(i)}$ = the set of indices of the components of $x^{(i)}$, $i = 1, 2$

$n \in \mathcal{N}^{(1)}$ if $n \notin \mathcal{N}'$ or $n \in \mathcal{N}'$ and $\bar{x}_n > 0$

$n \in \mathcal{N}^{(2)}$ if $n \in \mathcal{N}'$ and $\bar{x}_n = 0$

for a given $\bar{x} \in O_{f_0}$ and either component may be empty.

Note 1. When a vector a is partitioned into two subvectors, say

$$a = \{a^*, a^{**}\}$$

and we say that a^* (or a^{**}) is empty, this means that $a = a^{**}$ (or $a = a^*$).

Note 2. The above partitioning of the vector x obviously depends on the point $\bar{x}$ in O_{f_0} chosen. The same is true of the partitioning in (N.7) below and of various subsequent partitionings of x and g. It is understood that all these partitionings refer to the same choice of $\bar{x}$, and that $\bar{x}$, once chosen, remains fixed.

(N.7) $$g = \{g^{[1]}, g^{[2]}\}$$

where

$$g^{[1]}(\bar{x}) = 0, \quad g^{[2]}(\bar{x}) > 0$$

and either component may be empty.

(N.8) $$h(x) = 1 - g(x)$$

where 1 denotes the M-dimensional vector with 1's as components; $h^{[i]} = 1 - g^{[i]}$, $i = 1, 2$.

(N.9) $_{\eta_m}p_m(x) = 1 - [h_m(x)]^{1+\eta_m}, \quad m \in \mathcal{M}$.

(N.10) $\eta = \{\eta_1, \eta_2, \cdots, \eta_M\}$.

(N.11) $_\eta p(x) = \{\eta_1 p_1(x), \eta_2 p_2(x), \cdots, \eta p_M(x)\}$.

(N.12) $_\eta\phi(x, y) = f(x) + y'[\eta p(x)]$ (the "modified Lagrangian expression").

2.3.1. *A reformulation of Kuhn-Tucker theorem* 1. This slight generalization of theorem 1 (see [1], p. 484) is needed here because of the meaning of inequalities given in (N.1.4). [The possibility of this type of generalization is indicated in [1] (see pp. 491–492).][4]

We shall say that g satisfies the Constraint Qualification (C.Q.) at x, if the requirements of the definition in (A) of section 1.2 are satisfied with the inequalities (4), (5) in the same section interpreted in the sense of (N.1.4). $\phi(x, y)$ is given by (1) in 1.1. (It is immaterial whether g or $'g$ is used.)

THEOREM. *If f and g are differentiable, $\bar{x} \in O_{f_0}$ and g satisfy* C.Q. *at $\bar{x}$, then there exists a $\bar{y} \in Y$ such that*

$$\bar{y} \geqq 0;\ \bar{\phi}_y \cdot \bar{y} = 0;\ \bar{\phi}_y \cdot y \geqq 0 \textit{ for all } y \geqq 0;$$

$$\bar{x} \geqq 0;\ \bar{\phi}_x \cdot \bar{x} = 0;\ \bar{\phi}_x \cdot x \leqq 0 \textit{ for all } x \geqq 0.$$

[Note that, by virtue of the definitions in 2.2.2, this means that $\bar{\phi}_{y_m} \geqq 0$ if $m \in \mathcal{M}'$, $\bar{\phi}_{y_m} = 0$ if $m \notin \mathcal{M}'$, $\bar{\phi}_{x_n} \leqq 0$ if $n \in \mathcal{N}'$, $\bar{\phi}_{x_n} = 0$ if $n \notin \mathcal{N}'$. The other inequalities of the theorem are also to be interpreted in the sense of (N.1.4).]

[4] See also Hurwicz [9], pp. VIII — 2–6.

2.3.2. Theorem 1.

DEFINITION.[5] *An M-dimensional vector* $\eta = \{\eta_1, \eta_2, \cdots, \eta_M\}$ *is said to be acceptable if, for each* $m \in \mathcal{M}$, (1) $\eta_m \geqq 0$, *and* (2) η_m *is an even integer if* $h_m(\bar{x}) < -1$.

THEOREM 1. *If, for some* $\rho > 0$, $x \in S_\rho(\bar{x})$, $\bar{x} \in O_{fg}$, f *and* g *are differentiable, and* g *satisfies* C.Q. *at* $\bar{x}$, *then, for any acceptable* η, *there exists a vector* $\bar{y} = \bar{y}(\eta)$ *such that*

$$ {}_\eta\bar{\phi}_x \cdot x \leqq 0 \text{ for all } x \geqq 0 ; \tag{13}$$

$$ {}_\eta\bar{\phi}_x \cdot \bar{x} = 0 ; \tag{14}$$

$$ \bar{x} \geqq 0 ; \tag{15}$$

$$ {}_\eta\bar{\phi}_y \cdot y \geqq 0 \text{ for all } y \geqq 0 ; \tag{16}$$

$$ {}_\eta\bar{\phi}_y \cdot \bar{y} = 0 ; \tag{17}$$

$$ \bar{y} \geqq 0 . \tag{18}$$

The bar over ϕ denotes evaluation at $x = \bar{x}$, $y = \bar{y}(\eta)$.

Note that the relations (13)–(18) are necessary conditions for a nonnegative, in the sense of (N.1.4), saddle-point of ${}_\eta\phi(x, y)$ at $(\bar{x}, \bar{y})$. In particular, the relations (13)–(18) are satisfied if one selects $\bar{y} = \bar{y}(\eta)$ such that[6]

$$ (1 + \eta_m)\bar{y}_m(\eta) = \bar{y}_m(0) \text{ for all } m \in \mathcal{M} . \tag{19}$$

If the selection is made in accordance with (19), the equality

$$ {}_0\bar{\phi}_x = {}_\eta\bar{\phi}_x \tag{20}$$

will hold. Here ${}_0\phi(x, y)$ is ${}_\eta\phi(x, y)$ with $\eta = 0$; this is obviously the same as $\phi(x, y)$ in (1) of 1.1.

PROOF. For $\eta = 0$, the preceding theorem follows directly from the reformulated version of the Kuhn-Tucker theorem 1 given in 2.3.1. Thus there exists a vector

$$ \bar{y}(0) = \{\bar{y}_1(0), \bar{y}_2(0), \cdots, \bar{y}_M(0)\} \tag{21}$$

with the required properties.

Consider now the case $\eta \neq 0$. We shall show that $\bar{y}(\eta)$ defined by (19), that is, explicitly, by

$$ y_m(\eta) = \frac{1}{1 + \eta_m} \bar{y}_m(0) , \qquad m \in \mathcal{M} \tag{22}$$

[where $\bar{y}_m(0)$ is that of (21)], satisfies the relations (13)–(20).

We first observe that (22) yields

$$ (1 + \eta_m)\bar{y}_m(\eta)[h_m(\bar{x})]^{\eta_m} = \bar{y}_m(0) , \qquad m \in \mathcal{M} . \tag{23}$$

[When $h_m(\bar{x}) = 1$, (23) follows directly from (22). When $h_m(\bar{x}) \neq 1$, we have ${}_0\bar{\phi}_{y_m} = g_m(\bar{x}) > 0$, and hence, by (16)–(18), $\bar{y}_m(0) = 0$; (22) then yields $\bar{y}_m(\eta) = 0$ and (23) follows.]

Since

$$ {}_\eta\phi_{x_n} = f_{x_n} + \sum_{m=1}^{M} (1 + \eta_m) y_m(\eta) [h_m(x)]^{\eta_m} \frac{\partial g_m(x)}{\partial x_n} , \qquad n \in \mathcal{N} , \tag{24}$$

[5] In many applied problems, $h_m(x) \geqq 0$ for all m and all $x \geqq 0$. It was pointed out by Dr. Masao Fukuoka that, in the absence of such an assumption, the requirement of nonnegativity of the components of η is insufficient for the proof of the theorem.

[6] $\bar{y}_m(0) = \bar{y}$ in Kuhn-Tucker theorem 1 (see 2.3.1).

formula (23) implies

$$\eta\bar{\phi}_{x_n} = \bar{f}_{x_n} + \sum_{m=1}^{M} \bar{y}_m(0)\frac{\partial g_m(\bar{x})}{\partial x_n}, \qquad n \in \mathcal{N}. \tag{25}$$

Noting that the right member of (25) is identical with ${}_0\bar{\phi}_{x_n}$, we conclude that the relations (13)–(15) hold for all η with nonnegative components, since they are known to hold for $\eta = 0$.

Relation (16) is established by the fact that the right member of

$${}_\eta\bar{\phi}_{y_m} = {}_\eta p_m(\bar{x}) = 1 - [h_m(\bar{x})]^{1+\eta_m}, \qquad m \in \mathcal{M}, \tag{26}$$

is nonnegative for $m \in \mathcal{M}'$, zero for $m \notin \mathcal{M}'$ when η is acceptable (see the definition above) since, for any $m \in \mathcal{M}$, $h_m(\bar{x}) \leqq 1$, and, furthermore, ${}_\eta\bar{\phi}_{y_m} = 0$ if $m \notin \mathcal{M}'$, in which case $h_m(\bar{x}) = 1$.

Now suppose that, for some $m_0 \in \mathcal{M}$, ${}_\eta\bar{\phi}_{y_{m_0}} > 0$, that is, $h_{m_0}(\bar{x}) < 1$; then, by (16)–(18) for $\eta = 0$, $\bar{y}_{m_0}(0) = 0$; hence $\bar{y}_{m_0}(\eta) = 0$, and, therefore,

$$\bar{\phi}_{y_{m_0}} \cdot \bar{y}_{m_0}(\eta) = 0. \tag{27}$$

Since (27) clearly holds in the alternative case ${}_\eta\phi_{m_0} = 0$, (17) follows.

Finally, (18) holds because $\bar{y}_m(\eta)$ has the same sign as $\bar{y}_m(0)$ and the latter, by (18) for $\eta = 0$, is nonnegative if $m \in \mathcal{M}'$.

2.3.3. THEOREM 2. *Let, for some* $\rho > 0$, $x \in S_\rho(\bar{x})$, $\bar{x} \in O_{f_0}$, *such that* (13)–(20) *are satisfied. Then*

$${}_\eta\phi(\bar{x}, \bar{y}) \leqq {}_\eta\phi(\bar{x}, y) \text{ for all } y \geqq 0. \tag{28}$$

For we have

$${}_\eta\phi(\bar{x}, y) - {}_\eta\phi(\bar{x}, \bar{y}) = (y - \bar{y})\cdot{}_\eta\bar{p} = y\cdot{}_\eta\bar{p} \geqq 0 \text{ for } y \geqq 0 \tag{29}$$

where, since

$${}_\eta\bar{\phi}_y = {}_\eta\bar{p}, \tag{30}$$

the second equality follows from (17) and the inequality from (16).

2.3.4. *Notation.*

$$x^{(2)} = \{x^{(21)}, x^{(22)}\} \tag{N.13}$$

where

$${}_0\bar{\phi}_{x^{(21)}} = 0, \quad {}_0\bar{\phi}_{x^{(22)}} < 0$$

and either component may be empty.

$$x = \{x^I, x^{II}\} \tag{N.14}$$

where

$$\begin{aligned} x^I &= \{x^{(1)}, x^{(21)}\} \\ x^{II} &= x^{(22)}. \end{aligned} \tag{N.14.1}$$

(Either x^I or x^{II} may be empty.)

It should be noted that, by (13)–(15) and (N.13),

$$\begin{aligned} {}_0\bar{\phi}_{x^I} &= 0, \\ {}_0\bar{\phi}_{x^{II}} &< 0. \end{aligned} \tag{N.14.2}$$

2.3.5. *Definition of a regular constrained maximum.* In theorem 3 below we use the concept of a *regular* constrained maximum. The definition of such a maximum is given in the last part of this section. To state it, we must first formulate three regularity conditions denoted by R_1, R_2, R_3.

The first regularity condition R_1. Let $\bar{x}$ be a value maximizing the function $f(x)$ subject to $'g(x) \geqq 0$, $x \geqq 0$, and hence also subject to

$$g(x) \geqq 0 \qquad x \geqq 0 \tag{31}$$

where the inequalities are to be interpreted in the sense of (N.1.4).

From (N.6) and (N.7) it is clear that, for sufficiently small variations of x, the constraints

$$g^{[2]}(x) \geqq 0 \qquad x^{(1)} \geqq 0\,, \tag{32}$$

which are a part of (31), can be disregarded. Hence, at $\bar{x}$, $f(x)$ possesses a *local* maximum subject to

$$g^{[1]}(x) \geqq 0\,, \qquad x^{(2)} \geqq 0\,. \tag{33}$$

Let $g^{\dagger}$ be a subvector of $g^{[1]}$ such that $C_g = C_{\{g^{\dagger}, g^{[2]}\}}$ and write

$$g^{[1]} = \{g^{\dagger}, g^{\dagger\dagger}\}\,. \tag{34}$$

The components of $g^{\dagger\dagger}$ can be disregarded in the process of maximization, that is, $O_{f,g} = O_{f,\{g^{\dagger},g^{[2]}\}}$. If the Lagrangian multiplier vector $\bar{y}^{[1]}$ (corresponding to the constraints $g^{[1]}(x) \geqq 0$) is partitioned according to

$$\bar{y}^{[1]} = \{\bar{y}^{\dagger}, \bar{y}^{\dagger\dagger}\} \tag{35}$$

it is always possible to put

$$\bar{y}^{\dagger\dagger} = 0\,, \tag{36}$$

and this will be done in what follows.

Assuming that the constraints (33) are consistent, we may replace them by

$$g^{\dagger}(x) = 0 \qquad x^{(2)} \geqq 0\,. \tag{37}$$

The *first regularity condition* is

$$\text{rank}\,(\bar{g}^{\dagger}_{x^{(1)}}) = \dim g^{\dagger} = M^{\dagger}\,, \tag{R_1}$$

say.

Note 1. R_1 corresponds to the requirement of nondegeneracy in linear programming (see [4], p. 340).

Note 2. R_1 implies C.Q. (see appendix I).

The second regularity condition R_2. Since, by (N.7), (N.6), (N.14.1), and (34),

$$g^{\dagger}(\bar{x}) = 0 \qquad \bar{x}^{II} = 0\,, \tag{38}$$

it follows that, as a function of x^I, $f(x^I, \bar{x}^{II}) \equiv f(x^I, 0)$ has at $\bar{x}^I$ a local maximum subject to the constraints

$$(39) \qquad \begin{aligned} g^\dagger(x^I, \bar{x}^{II}) \equiv g^\dagger(x^I, 0) &= 0 \\ x^{(21)} &\geqq 0 \,. \end{aligned}$$

The corresponding Lagrangian expression becomes

$$(40) \qquad {}_0\phi^I(x^I, y^\dagger) = f(x^I, 0) + y^\dagger \cdot g^\dagger(x^I, 0) \,.$$

Using the reformulation of Kuhn-Tucker theorem 1, given in 2.3.1, we may assert the existence of a $\bar{y}^\dagger$ such that

$$(41) \qquad \bar{x}^I \geqq 0; \; {}_0\bar{\phi}^I_{x^I} = 0 \,;$$

$$(42) \qquad \bar{y}^\dagger \geqq 0; \; {}_0\bar{\phi}^I_{y^\dagger} = 0 \,.$$

It might happen that some components of $\bar{y}^\dagger$ vanish. Write $y^\dagger = \{y^*, y^0\}$ where every component of $\bar{y}^*$ is different from zero and

$$(43) \qquad \bar{y}^0 = 0 \,.$$

Let $g^\dagger$ be correspondingly partitioned as

$$(44) \qquad g^\dagger = \{g^*, g^0\} \,.$$

Now suppose that ${}_0\phi^I(x^I, y^\dagger)$ has a nonnegative saddle-point at $(\bar{x}^I, \bar{y}^\dagger)$. By theorem 3 in Kuhn-Tucker, a sufficient condition for this is that f and g be both concave. One can then easily verify that

$$(45) \qquad {}_0\phi^I(x^I, y^*) \equiv f(x^I, 0) + y^* \cdot g^*(x^I, 0)$$

has a nonnegative saddle-point at $(\bar{x}^I, \bar{y}^*)$.

But then $\bar{x}^I$ maximizes $f(x^I, 0)$ subject to $g^*(x^I, 0) \geqq 0$ and $x^{(21)} \geqq 0$. Hence in this case the components of g^0 could have been disregarded in the original maximation problem ($O_{f,g} = O_{f,\{g^*,g[2]\}}$).

However, complications might arise if ${}_0\phi^I(x^I, y^\dagger)$ did not have a nonnegative saddle-point at $(\bar{x}^I, \bar{y}^\dagger)$. To take care of this case, one might require that

(46) g^0 is empty unless ${}_0\phi^I(x^I, y^\dagger)$ has a local nonnegative saddle-point at $(\bar{x}^I, \bar{y}^\dagger)$.

However, to simplify matters we shall impose the seemingly[7] stronger condition

(47) g^0 is empty.

It follows that

$$(48) \qquad M^* \equiv \dim g^* = \dim g^\dagger \equiv M^\dagger \,.$$

Let $\mathcal{M}^*$ [$= \mathcal{M}^\dagger$ by (47)] denote the set of indices of g^*. Clearly, for $m \in \mathcal{M}^* \cap (\mathcal{M} \sim \mathcal{M}')$, we may have $\bar{y}_m < 0$.

Now suppose the preceding reasoning had been carried out in terms of $'g$ instead of g. Nothing would be changed, except, possibly, the signs of some components of the Lagrangian multiplier, to be denoted by $'\bar{y}$.

That is, we would have $'\bar{y}_m > 0$ for $m \in \mathcal{M}^* \cap \mathcal{M}'$ and $'\bar{y}_m > 0$ or $'\bar{y}_m < 0$ for

[7] See section 2.3.7.

$m \in \mathcal{M}^* \cap (\mathcal{M} \sim \mathcal{M}')$. Let $\mathcal{M}^{\cdot}$ be defined by the relation $m \in \mathcal{M}^{\cdot}$ if and only if $m \in \mathcal{M}^* \cap (\mathcal{M} \sim \mathcal{M}')$ and $'\bar{y}_m < 0$. Then, it is clear from (N.2.2) that we may put

$$\begin{aligned} \bar{y}_m &= {}'\bar{y}_m \text{ for } m \in \mathcal{M} \sim \mathcal{M}^{\cdot} \\ \bar{y}_m &= -{}'\bar{y}_m \text{ for } m \in \mathcal{M}^{\cdot} , \end{aligned} \tag{49}$$

so that $\bar{y}_m > 0$ for all $m \in \mathcal{M}^*$.

Hence, *without loss of generality* [as compared with (47)] condition (47) may be restated as the *second regularity condition,*

$$\begin{aligned} &g^0 \text{ is empty and} \\ &\bar{y}_m > 0 \text{ for all } M \in \mathcal{M}^* . \end{aligned} \tag{R$_2$}$$

The first regularity condition then implies

$$\operatorname{rank} (\bar{g}^*_{x^{(1)}}) = M^* \tag{50}$$

where

$$M^* = \dim g^* . \tag{51}$$

The third regularity condition R_3. When the first two regularity conditions are satisfied, second derivatives are continuous, and x^I is nonempty, it is possible to show (see appendix II) that a certain quadratic form is nonpositive when some of the variables are restricted in sign. The third regularity condition is a strengthening of (71) requiring that the quadratic form in question be negative under the same restrictions. This condition, analogous to that used by Samuelson (see [5], p. 358) makes it possible to avoid going beyond second order terms in the expansions used.

The third regularity condition is formulated in terms of a function $q(t)$ of a new variable vector

$$t = \{t^*, t^{**}\} \tag{52}$$

which is obtained by a transformation of coordinates from x^I after the latter has been partitioned so that

$$x^I = \{x^*, x^{**}\} , \tag{53}$$

where x^* is a subvector of $x^{(1)}$.

We shall (a) define x^* and x^{**}; (b) write down the transformation defining $\{t^*, t^{**}\}$ in terms of $\{x^*, x^{**}\}$; (c) define $q(t)$; and (d) formulate the third regularity condition.

In the remainder of this section it is assumed that R_1 holds; it is also assumed that x^I is not empty.

First case: $M^* = 0$. Write

$$t = t^{**} = x^{**} = x^I , \tag{54}$$

so that, by (52) and (53), x^* and t^* are empty, and define

$$q(t) = f(x^I, \bar{x}^{II}) = f(t^{**}, 0) . \tag{55}$$

The third regularity condition for this case is formulated in R_3 below.

Second case: $M^* > 0$. (a) *The definition of* x^*. From R_1 it follows that there exists a (nonempty) M^*-dimensional subvector x^* of $x^{(1)}$ such that

$$\bar{g}^*_{x^*} \text{ is an } M^* \text{ by } M^* \ (M^* \geqq 1) \text{ nonsingular matrix} . \tag{56}$$

We then define x^{**} by (53) and $x^{(12)}$ by

$$x^{(1)} = \{x^*, x^{(12)}\}\,. \tag{57}$$

Clearly

$$x^{**} = \{x^{(12)}, x^{(21)}\}\,. \tag{58}$$

(b) *The transformation from* x^I *to* t. Let

$$h^* = 1 - g^* \tag{59}$$

where 1 is the M^*-dimensional vector with (scalar) 1's as components. $t = \{t^*, t^{**}\}$ is then defined by the transformation

$$t^* = h^*(x^*, x^{**}, \bar{x}^{II}) \tag{60}$$

$$t^{**} = x^{**}\,. \tag{61}$$

We also partition t^{**} by

$$t^{**} = \{t^{(12)}, t^{(21)}\} \tag{62}$$

where

$$\begin{aligned} t^{(12)} &= x^{(12)}\,, \\ t^{(21)} &= x^{(21)}\,. \end{aligned} \tag{63}$$

This is obviously consistent with (57) and (61).

(c) *The definition of* $q(t)$. By (59), the Jacobian H of the transformation (58)–(59) is

$$H = \begin{pmatrix} h^*_{x^*} & h^*_{x^{**}} \\ 0 & I \end{pmatrix} = -\begin{pmatrix} g^*_{x^*} & g^*_{x^{**}} \\ 0 & -I \end{pmatrix}, \tag{64}$$

so that, by (56),

$$|\bar{H}| = -\,|-\bar{g}^*_{x^*}| \neq 0\,, \tag{65}$$

that is,

$$\bar{H} \text{ is nonsingular}\,. \tag{66}$$

Hence, locally, (60)–(61) can be solved for x^I in terms of t; we may write this solution as

$$x^I = r(t) \tag{67}$$

where

$$r = \{r^*, r^{**}\} \tag{68}$$

and

$$x^* = r^*(t),\; x^{**} = r^{**}(t) = t^{**}\,. \tag{69}$$

The function $q(t)$ is now defined as $f(x)$ evaluated at $x^{II} = \bar{x}^{II}$ and with x^I expressed in terms of t, that is,

$$q(t) = f[r(t), \bar{x}^{II}] \equiv f[r^*(t^*, t^{**}), t^{**}, 0]\,. \tag{70}$$

The statement of the third regularity condition. We have now defined $q(t)$ for all M^* provided the first regularity condition R_1 is satisfied and x^I is nonempty. It is shown in appendix II that, assuming R_1, R_2, and the continuity of the second derivatives, unless x^{**} is empty, there exists $\rho > 0$ such that, for all $t^{**} \in S_\rho(\bar{x}^{**})$,

$$(t^{**} - \bar{x}^{**})' \, \bar{q}_{t^{**}t^{**}} \, (t^{**} - \bar{x}^{**}) \leqq 0 \,, \qquad \text{if } t^{(21)} \geqq 0. \tag{71}$$

The *third regularity condition* is a strengthening of the preceding inequality. It states that

(R_3) (a) x^{**} is empty or

(b) there exists $\rho > 0$ such that, for all $t^{**} \in S_\rho(\bar{x}^{**})$, $(t^{**} - \bar{x}^{**})' \, \bar{q}_{t^{**}t^{**}}(t^{**} - \bar{x}^{**}) < 0$ if $t^{(21)} \geqq 0$ and $t^{**} \neq \bar{x}^{**}$.

Note. The situation covered by (a) of R_3 is of importance since it permits the treatment of a large class of cases where f and g are linear.

DEFINITION. *$f(x)$ is said to have a regular maximum at $\bar{x}$ subject to $g(x) \geqq 0$, $x \geqq 0$, if the three regularity conditions* R_1, R_2, R_3 *are satisfied at $\bar{x}$ and $\bar{x} \in O_{fg}$.*

2.3.6. THEOREM 3. *If, for some $\rho > 0$, $x \in S_\rho(\bar{x})$, $\bar{x}$ a regular maximum*[8] *of $f(x)$ subject to $g(x) \geqq 0$ and $x \geqq 0$, f and g are differentiable (with regard to x), and furthermore, when x^I is nonempty, have continuous second order derivatives with regard to x^I, then, for all acceptable*[9] *η sufficiently large in each component,*

$$x^I \text{ is empty}\,, \tag{72}$$

or

$$(x^I - \bar{x}^I)' {}_\eta\bar{\phi}_{x^Ix^I}(x^I - \bar{x}^I) < 0 \text{ if } x^{(21)} \geqq 0,\ x^I \neq \bar{x}^I\,, \tag{73}$$

and for some $\rho' > 0$, and all $x \in S_{\rho'}(\bar{x})$ such that $x \geqq 0$, $x \neq \bar{x}$,

$${}_\eta\phi[x, \bar{y}(\eta)] < {}_\eta\phi[\bar{x}, \bar{y}(\eta)] \tag{74}$$

where ${}_\eta\phi$ and $\bar{y}(\eta)$ are defined as in theorem 1.

Note.[10] Theorem 3 is valid for f, g linear if x^{**} is empty (regardless of whether x^* is empty), provided the first two regularity conditions hold. However, if both x^* and x^{**} are empty, x^I is empty, and the theorem follows from the first case considered below. If x^{**} is empty while x^* is nonempty, use the first two cases below together with (90) (since g^* is nonempty and t^{**} is empty). Note that x^{**} is empty at the basic solutions of a linear programming problem.

2.3.7. *Proof of theorem* 3. First it is shown that (72) or (73) implies (74). Then it is shown that (72) or (73) is true.

It can be seen that if theorem 3 is established for the case of $\{g^{\dagger\dagger}, g^0\}$ empty, then theorem 3 is also true if (i) $g^{\dagger\dagger}$ is not empty, and/or (ii) g^0 is not empty but ${}_0\phi^I(x^I, y^\dagger)$ has a nonnegative saddle-point at $(\bar{x}^I, \bar{y}^\dagger)$, since in either case $\bar{x}$ remains unchanged and the additional terms in the modified Lagrangian expression vanish at $\bar{y}$ [compare equations (36) and (43)].

Hence, with no loss of generality, we may henceforth assume $\{g^{\dagger\dagger}, g^0\}$ to be empty, that is,

$$g^{[1]} = g^*\,. \tag{75}$$

We now show that (72) or (73) implies (74), that is, that in a sufficiently small neighborhood, if (72) or (73) is assumed to be valid and the inequalities $x \geqq 0$, $x \neq \bar{x}$, hold,

[8] The term "regular maximum" is defined at the end of section 2.3.5.

[9] The term "acceptable" is defined at the beginning of section 2.3.2.

[10] The desirability of explicit treatment of the linear case was emphasized by Dr. Masao Fukuoka.

the conclusion of (74) follows. We write ϕ instead of ${}_\eta\phi$ throughout. Also (72) or (73), $x \geqq 0$, $x \neq \bar{x}$, is assumed.

Let

$$\xi = x - \bar{x} \tag{76}$$

$$\xi^I = x^i - \bar{x}^i\,, \qquad i = I, II. \tag{77}$$

First case: $\xi^{II} \neq 0$. By (20) and (N.14.2),

$$\bar{\phi}_x \cdot \xi = \bar{\phi}_{x^I} \cdot \xi^I + \bar{\phi}_{x^{II}} \cdot \xi^{II} < 0\,. \tag{78}$$

But then the conclusion of (74) follows from the well-known "Fréchet" property of differentials[11] which, as applied to the present case, states that, given any $\sigma > 0$, there exists an $\epsilon > 0$ such that,

$$\left| \frac{1}{|\xi|} [\phi(x, \bar{y}) - \phi(\bar{x}, \bar{y}) - \bar{\phi}_x \cdot \xi] \right| < \sigma \tag{79}$$

if $|\xi| < \epsilon$.

Choose

$$\sigma = -\frac{1}{|\xi|} \bar{\phi}_x \cdot \xi \tag{80}$$

which is positive by (78). Then, for a sufficiently small $|\xi|$, we have by (79)

$$\left| \frac{1}{|\xi|} [\phi(x, \bar{y}) - \phi(\bar{x}, \bar{y}) + \sigma] \right| < \sigma \tag{81}$$

which implies

$$\frac{1}{|\xi|} [\phi(x, \bar{y}) - \phi(\bar{x}, \bar{y})] < 0 \tag{82}$$

and hence the conclusion of (74).

If x^I is empty, this completes the proof of the theorem 3, since $x \neq \bar{x}$ then implies $\xi^{II} \neq 0$. If x^I is not empty, we must consider the

Second case: $\xi^{II} = 0$. Since it is assumed that $x \neq \bar{x}$, $\xi^{II} = 0$ implies

$$\xi^I \neq 0\,. \tag{83}$$

In virtue of the existence of the second derivatives of ϕ with regard to x^I (by definition of ϕ, and the assumptions concerning the second derivatives of f and g with regard to x^I) we have, by Taylor's theorem,

$$\phi(x, \bar{y}) - \phi(\bar{x}, \bar{y}) = \bar{\phi}_{x^I} \cdot \xi^I + \tfrac{1}{2}(\xi^I)' \tilde{\phi}_{x^I x^I} \xi^I\,, \tag{84}$$

where $\tilde{\phi}_{x^I x^I}$ denotes $\phi_{x^I x^I}$ evaluated at $x = \tilde{x}$, $\tilde{x} = \bar{x} + \theta\xi$, $0 < \theta < 1$. It now suffices to note that $(\xi^I)' \phi_{x^I x^I} \xi^I$ is negative at $\bar{x}$ [since (72) or (73) is assumed to hold and its hypotheses are satisfied] and continuous in the neighborhood of $\bar{x}$ (by the hypotheses of the theorem concerning the second derivatives of f and g), so that, for a sufficiently small $|\xi^I|$, $(\xi^I)' \tilde{\phi}_{x^I x^I} \xi^I < 0$. Since $\bar{\phi}_{x^I} \cdot \xi^I = 0$ by (N.14.2), (74) follows.

We now show that (72) holds if x^I is nonempty.

First case: g^* empty. By equation (75), $g^{[1]}$ is also empty. Hence, by (13)–(15) in theorem 1,

$$\bar{y}(\eta) = \bar{y}^{[2]}(0) = 0 \tag{85}$$

[11] See Hille [10], p. 72, definition 4.3.4, equation (iii).

and, using (N.12),

$$(86) \qquad {}_\eta\phi[x, \bar{y}(\eta)] = f(x) .$$

Since g^* is empty, we have $M^* = 0$, and, therefore, the definition (55) of q applies, so that (since x^* is empty but x^I is not) t^{**} is not empty and

$$(87) \qquad \bar{q}_{t^{**}t^{**}} = \bar{f}_{x^{**}x^{**}} = {}_\eta\bar{\phi}_{x^Ix^I} .$$

Equations (86) and (87), together with the third regularity condition R_3, yield (73) for a sufficiently small neighborhood of $\bar{x}$.

Second case: g^* nonempty. Write

$$(88) \qquad \psi(t, y) = \phi[r(t), \bar{x}^{II}, y]$$

where $r(t)$ is defined in (67). (Where it is desired to indicate the dependence of ψ on η, we may write ${}_\eta\psi$ instead of ψ.)

Then, by (66), that is, R_1, we have

$$(89) \qquad \bar{\Psi}_{tt} = \psi_{tt}\big|_{t=\bar{t}} = (\bar{H}^{-1})'{}_\eta\bar{\phi}_{x^Ix^I}\bar{H}^{-1}, \bar{t} = \{h^*(\bar{x}), \bar{x}^{**}\} ,$$

since ${}_\eta\bar{\phi}_{x^I} = {}_0\bar{\phi}_{x^I} = 0$ by (20) and (N.14.2).

We shall now show that (73) is implied by

$$(90) \qquad \tau'\bar{\Psi}_{tt}\tau < 0, \text{ if } \tau^{(21)} \geqq 0, \text{ and } \tau \neq 0$$

where the partitioning of τ corresponds to that of t. We show later that (90) holds.

To see that (90) implies (73), let x^I satisfy the inequalities $x^{(21)} \geqq 0$, $x^I \neq \bar{x}^I$. Choose

$$(91) \qquad \begin{pmatrix} \tau^* \\ \tau^{**} \end{pmatrix} = \tau = \bar{H}\,(x^I - \bar{x}^I) = \begin{pmatrix} \bar{h}^*_{x^*} & \bar{h}^*_{x^{**}} \\ 0 & I \end{pmatrix} \begin{pmatrix} x^* - \bar{x}^* \\ x^{**} - \bar{x}^{**} \end{pmatrix}.$$

Since, by (66), $\bar{H}$ is nonsingular, $x^I \neq \bar{x}^I$ implies $\tau \neq 0$. Also, (91) yields

$$(92) \qquad \tau^{**} = x^{**} - \bar{x}^{**} ,$$

hence, in particular,

$$(93) \qquad \tau^{(21)} = x^{(21)} - \bar{x}^{(21)} .$$

But

$$(94) \qquad \bar{x}^{(21)} = 0 ,$$

since $x^{(21)}$ is a component of $x^{(2)}$ by (N.13), and $\bar{x}^{(2)} = 0$ by (N.6). Hence

$$(95) \qquad \tau^{(21)} = x^{(21)}$$

and thus $x^{(21)} \geqq 0$ implies $\tau^{(21)} \geqq 0$.

Having shown that the hypotheses of (73) imply those of (90), we see that the hypotheses of (73), together with the validity of the assertion in (90), yield

$$(96) \qquad \tau'\bar{\Psi}_{tt}\tau < 0 .$$

But, using in succession (91), (89), and simplifying, we have

$$(97) \qquad \begin{aligned} \tau'\psi_{tt}\tau &= (x^I - \bar{x}^I)'\bar{H}'\bar{\Psi}_{tt}\bar{H}(x^I - \bar{x}^I) \\ &= (x^I - \bar{x}^I)'\bar{H}'(\bar{H}^{-1})'{}_\eta\phi_{x^Ix^I}\bar{H}^{-1}\bar{H}(x^I - \bar{x}^I) \\ &= (x^I - \bar{x}^I)'{}_\eta\bar{\phi}_{x^Ix^I}(x^I - \bar{x}^I) . \end{aligned}$$

Formulas (96) and (97) yield the conclusion of (73). Thus it has been established that (90) implies (73). It remains to be shown that (90) is valid. It is convenient to write $\bar{\psi}_{tt}$ in the partitioned form

$$\bar{\psi}_{tt} = \begin{pmatrix} \bar{\psi}_{t^*t^*} & \bar{\psi}_{t^*t^{**}} \\ \bar{\psi}_{t^{**}t^*} & \bar{\psi}_{t^{**}t^{**}} \end{pmatrix} \equiv \begin{pmatrix} A & B \\ B' & C \end{pmatrix} \tag{98}$$

where t^{**} may be empty; t^* is assumed nonempty, since the case of t^* empty was treated earlier.

It will now be shown that A, that is, $\bar{\psi}_{t^*t^*}$ [compare (98)], which depends on η, can be made negative definite by a suitable choice of η.

Recalling that $\mathcal{M}^*$ denotes the set of indices of the components of g^*, and using (N.9) and (60), we see that, for $m \in \mathcal{M}^*$,

$$_{\eta_m}p_m(x) = 1 - t_m^{1+\eta_m} \tag{99}$$

where t_m is a component of t^*.

Since, by theorem 1 and equation (75),

$$\bar{y}_m(\eta) = 0 \text{ for } m \in \mathcal{M} \sim \mathcal{M}^*, \tag{100}$$

we have, from the definitions of ψ, q, and $_\eta\phi$ [equations (88), (70), and (N.12), respectively], and the preceding relations (99) and (100), the equality

$$\psi\,[t, \bar{y}\,(\eta)] = q\,(t) + \sum_{m \in \mathcal{M}^*} [\bar{y}_m\,(\eta)]\,(1 - t_m^{1+\eta_m})\,. \tag{101}$$

Writing

$$F = \bar{q}_{t^*t^*}\,, \tag{102}$$

we have, from (101) and the definition of A that

$$A = F - D\,, \tag{103}$$

where $D = ||d_{m,m'}||$, $m \in \mathcal{M}^*$, $m' \in \mathcal{M}^*$, is a diagonal matrix [that is, $d_{m,m'} = 0$ for $m \neq m'$] with

$$d_{m,m} = [\bar{y}_m(\eta)](1 + \eta_m)\eta_m = \bar{y}_m(0)\eta_m\,, \qquad m \in \mathcal{M}^*, \tag{104}$$

where the second equality follows from (19).

Let λ denote the largest characteristic root of F. Since, by the second regularity condition R_2, $\bar{y}_m(0) > 0$ if $m \in \mathcal{M}^*$, we may choose η_m^0, for each $m \in \mathcal{M}^*$, to be a positive even integer satisfying

$$\eta_m^0 > \lambda/\bar{y}_m(0)\,, \tag{105}$$

so that

$$\min_{m \in \mathcal{M}^*} d_{m,m} > \lambda \tag{106}$$

for all acceptable $\eta_m \geqq \eta_m^0$.

Then, for any $t^* \neq 0$, and each acceptable $\eta_m \geqq \eta_m^0$, we have

$$\begin{aligned} t^{*\prime}Ft^* \leqq \lambda t^{*\prime}t^* \equiv \lambda \sum_{m \in \mathcal{M}^*} t_m^2 < \sum_{m \in \mathcal{M}^*} d_{m,m}\, t_m^2 \\ \equiv t^{*\prime}Dt^*\,, \end{aligned} \tag{107}$$

that is, $t^* \neq 0$ implies $t^{*\prime}(F - D)t^* < 0$ for all sufficiently large acceptable η, or A is negative definite for all sufficiently large acceptable η.

This suffices to establish (90) and, therefore, (73) if t^{**} is empty.

Now assume t^{**} not empty. Write

$$P = \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix} \tag{108}$$

and

$$_\eta\Omega = P'{}_\eta\bar{\Psi}_{tt}P\,. \tag{109}$$

Then methods used to show that (90) implies (73) can be used to show that

$$w'{}_\eta\Omega w < 0 \text{ for } w \neq 0,\ w^{(21)} \geqq 0 \tag{110}$$

implies (90). This is because

$$P^{-1} = \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix} \tag{111}$$

and, like its analogue $\bar{H}$, performs an identity transformation on t^{**}, so that the condition $t^{(21)} \geqq 0$ is transformed into the condition $w^{(21)} \geqq 0$. It remains to establish (110). Now from (109), (108), and (98), we have

$$_\eta\Omega = \begin{pmatrix} A & 0 \\ 0 & C - B'A^{-1}B \end{pmatrix}, \tag{112}$$

so that $w'_\eta\Omega w = w^{*\prime}Aw^* + w^{**\prime}\,(C - B'A^{-1}B)w^{**}$.

Now, we may take A as negative definite, and hence, to establish (110), it will suffice to show that

$$Q \equiv w^{**\prime}(C - B'A^{-1}B)w^{**} < 0, \text{ if } w^{**} \neq 0,\ w^{(21)} \geqq 0\,. \tag{113}$$

Before doing so, we shall obtain an auxiliary result.

It will now be shown that the norm of A^{-1} can be made arbitrarily small by choosing η sufficiently large. It does not matter which of the many norms is used (see Bowker [6]). Note that, denoting by $N(X)$ the norm of the matrix X, we have $N(A + B) \leqq N(A) + N(B)$, $N(AB) \leqq N(A)N(B)$; if all the elements of a matrix approach 0, so does its norm. If I denotes the identity matrix, $N(I) = 1$.

D^{-1} is a diagonal matrix whose nonzero elements approach zero for η large; hence, the same is true of $D^{-1}F$. Therefore, η can be chosen sufficiently large so that,

$$I - D^{-1}F \text{ is nonsingular}, \tag{114}$$

and

$$N(D^{-1}F) < 1\,. \tag{115}$$

Following Waugh (see p. 148, [7]), we use the identity, valid because of (114),

$$(I - D^{-1}F)^{-1} = I + (I - D^{-1}F)^{-1}D^{-1}F\,, \tag{116}$$

and the properties of the norm to derive the relation,

$$N[(I - D^{-1}F)^{-1}] \leqq 1 + N[(I - D^{-1}F)^{-1}]N(D^{-1}F)\,. \tag{117}$$

From (117) and (115), it follows that,

$$N[(D^{-1}F - I)^{-1}] \leqq \frac{1}{1 - N(D^{-1}F)}\,. \tag{118}$$

Since $A = F - D = D(D^{-1}F - I)$, it follows that $A^{-1} = (D^{-1}F - I)^{-1}D^{-1}$, and hence

$$(119)\qquad N(A^{-1}) \leqq N(D^{-1})\,N[(D^{-1}F - I)^{-1}] \leqq \frac{N(D^{-1})}{1 - N(D^{-1}F)}$$

which can be made arbitrarily small for η large.

Consider now the quadratic form Q in (113). We have shown, using (101), that

$$(120)\qquad C = \bar{\psi}_{t^{**}t^{**}} = \bar{q}_{t^{**}t^{**}}\,.$$

Hence the third regularity condition, R_3, implies

$$(121)\qquad w^{**\prime}Cw^{**} < 0 \text{ if } w^{**} \neq 0,\ w^{(21)} \geqq 0\,.$$

As shown earlier $N(B'A^{-1}B) \leqq N(B')N(A^{-1})N(B) = N(A^{-1})[N(B)]^2$ can be made arbitrarily small by choosing a large enough η. Now

$$(122)\qquad |w^{**\prime}B'A^{-1}Bw^{**}| \leqq N(B'A^{-1}B)w^{**\prime}w^{**}\,,$$

since the characteristic roots of a matrix are bounded in absolute value by its norm.

Also, denoting by μ the maximum of $w^{**\prime}Cw^{**}$ subject to $w^{**\prime}w^{**} = 1$, $w^{(21)} \geqq 0$, we have

$$(123)\qquad w^{**\prime}Cw^{**} \leqq \mu w^{**\prime}w^{**}$$

and, by (121), $\mu < 0$. With the aid of (122),

$$(124)\qquad Q < [\mu + N(B'A^{-1}B)]w^{**\prime}w^{**} \text{ if } w^{**} \neq 0,\ w^{(21)} \geqq 0\,.$$

By choosing η sufficiently large, so that

$$(125)\qquad \mu + N(B'A^{-1}B) < 0\,,$$

we establish (113), which, in turn, yields (110), (90), (73), and hence theorem 3.

APPENDIX I[12]

Let the first regularity condition R_1 hold. Consider $\bar{x}$ such that,

$$(126)\qquad g^{[1]}(\bar{x}) = 0,\ g^{[2]}(\bar{x}) > 0,\ \bar{x} \geqq 0\,,$$

and x such that,

$$(127)\qquad \bar{g}_x^{[1]}(x - \bar{x}) \geqq 0,\ x^{(2)} - \bar{x}^{(2)} \geqq 0\,,$$

where all inequalities are to be interpreted in the sense of (N.1.4). Define now the function $g^{\#}$ of $\bar{x}$ by

$$(128)\qquad g^{\#}(x) = \{g^{\dagger}(x),\ x^{**},\ x^{II}\}$$

where x^{**} is defined by (58). Notice that assuming g^0 to be empty as in (47), $g^{\#}$, like x, has N dimensions.

It follows that

$$(129)\qquad g_x^{\#} = \begin{pmatrix} g_{x^*}^{\dagger} & g_{x^{**}}^{\dagger} & g_{x^{II}}^{\dagger} \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix}$$

and hence

$$(130)\qquad |\bar{g}_x^{\#}| = |\bar{g}_{x^*}^{\dagger}| \neq 0\,.$$

[12] This appendix parallels lemma 76.1 in Bliss [11].

Consider now the relation which associates with a real number α the values $\bar{\bar{x}}$ of x for which the equation

$$g^{\#}(\bar{\bar{x}}) = g^{\#}(\bar{x}) + \alpha \bar{g}_x^{\#}(x - \bar{x}) \tag{131}$$

is satisfied. In virtue of the implicit function theorem, for sufficiently small values of α(131) defines $\bar{\bar{x}}$ as a (single-valued) differentiable function of α, say

$$\bar{\bar{x}} = \psi_1(\alpha) , \tag{132}$$

such that

$$\psi_1(0) = \bar{x} . \tag{133}$$

Differentiating (131) with respect to α and setting $\alpha = 0$, we have

$$\bar{g}_x^{\#}\psi_1'(0) = \bar{g}_x^{\#}(x - \bar{x}) \tag{134}$$

and hence, because of (130),

$$\psi_1'(0) = x - \bar{x} . \tag{135}$$

We shall now show that

$$\psi_1(\alpha) \in C_g \text{ for } \alpha \geqq 0, \ \alpha \text{ sufficiently small} . \tag{136}$$

By (131), (126), and (127)

$$g^{\dagger}(\bar{\bar{x}}) = \alpha \bar{g}_x^{\dagger}(x - \bar{x}) \geqq 0 \text{ for } \alpha \geqq 0 . \tag{137}$$

It follows that

$$g^{[1]}(\bar{\bar{x}}) \geqq 0 \text{ for } \alpha \geqq 0 , \tag{138}$$

which together with

$$g^{[2]}[\psi_1(\alpha)] \geqq 0 \text{ for } \alpha \text{ sufficiently small} , \tag{139}$$

yields

$$g[\psi_1(\alpha)] \geqq 0 \text{ for } \alpha \geqq 0 \text{ sufficiently small} . \tag{140}$$

Now, since x^* is a subvector of $x^{(1)}$, $x^{(2)}$ is a subvector of $\{x^{**}, x^{II}\}$, hence (127) and (131) imply

$$\bar{\bar{x}}^{(2)} = \psi_1^{(2)}(\alpha) = \bar{x}^{(2)} + \alpha(x^{(2)} - \bar{x}^{(2)}) \geqq 0 \text{ for } \alpha \geqq 0 \tag{141}$$

which, together with

$$\bar{\bar{x}}^{(1)} = \psi_1^{(1)}(\alpha) \geqq 0 \text{ for } \alpha \text{ sufficiently small, yields} \tag{142}$$

$$\psi_1(\alpha) \geqq 0 \text{ for } \alpha \geqq 0, \ \alpha \text{ sufficiently small} . \tag{143}$$

In turn, (140) and (143) yield (136).

Now let us interpret "α sufficiently small" as $0 \leqq \alpha \leqq \lambda$ where $\lambda > 0$ and define the function ψ by

$$\psi(\theta) = \psi_1(\lambda\theta) \text{ for all } 0 \leqq \theta \leqq 1 . \tag{144}$$

Then

$$\begin{aligned} &\psi(0) = \bar{x} , && \\ &\psi'(0) = \lambda\psi_1'(0) = \lambda(x - \bar{x}) , && \lambda > 0 , \\ &\psi(\theta) \in C_g , && 0 \leqq \theta \leqq 1 . \end{aligned} \tag{145}$$

Since (145) are precisely the requirements of C.Q., it has been shown that R_1 implies C.Q.

APPENDIX II

We shall now show that, if the first two regularity conditions hold and if in a neighborhood of $\bar{x}$, f and g are assumed to possess continuous derivatives of second order with regard to x^I, then (71) is valid.

Let x^{**} be nonempty. Then, writing

$$\bar{t}^* = h^*(\bar{x}) = 1 \qquad \text{(a vector of 1's)}, \tag{146}$$

$$\bar{t}^{**} = \bar{x}^{**}, \tag{147}$$

we have, using Taylor's theorem,

$$q(\bar{t}^*, t^{**}) - q(\bar{t}^*, \bar{t}^{**}) = \bar{q}_{t^{**}} \cdot (t^{**} - \bar{t}^{**}) + \tfrac{1}{2}(t^{**} - \bar{t}^{**})' \tilde{q}_{t^{**}t^{**}}(t^{**} - \bar{t}^{**}), \tag{148}$$

where "$-$" over a symbol denotes the evaluation at $t = \bar{t}$, while "$\sim$" over a symbol denotes evaluation at $t = \tilde{t}$, where $\tilde{t} = \bar{t} + \theta\,(t^{**} - \bar{t}^{**})$, $0 < \theta < 1$. Now suppose it has been shown that (a) $q(\bar{t}^*, t^{**})$ has, as a function of t^{**}, subject to the constraint $t^{(21)} \geqq 0$, a local maximum at $t^{**} = \bar{t}^{**}$, and (b) $\bar{q}_{t^{**}} = 0$. From (a) it follows that, in a sufficiently small neighborhood, the left member of (148) is nonpositive if $t^{(21)} \geqq 0$. But then, using (b), we see that the quadratic form in the right member of (148) is nonpositive. Since, by hypothesis, $q_{t^{**}t^{**}}$ is a continuous function of t^{**}, we have, for $t^{(21)} \geqq 0$, and in a sufficiently small neighborhood of $\bar{t}$,

$$(t^{**} - \bar{t}^{**})\bar{q}_{t^{**}t^{**}}(t^{**} - \bar{t}^{**}) \geqq 0 \tag{149}$$

which is the desired result (71). Hence it remains to prove (a) and (b).

(a) *$q(\bar{t}^*, t^{**})$ has, as a function of t^{**}, subject to $t^{(21)} \geqq 0$, a local maximum at $t^{**} = \bar{t}^{**}$.*

It follows from the remarks at the beginning of the discussion of the second regularity condition that $f(x^I, 0)$, as a function of x^I, has a local maximum at $x^I = \bar{x}^I$, subject to the constraints

$$g^{\dagger}(x^I, 0) = 0, \; x^{(21)} \geqq 0. \tag{150}$$

Hence, subject to the same constraints, $q(t)$ has a local maximum at $\bar{t}$. Now we must distinguish the two ways in which the "milder" (46) second regularity condition R_2 may be satisfied.

First way: ${}_0\phi^I(x^I, y^{\dagger})$ has a nonnegative saddle-point at $(\bar{x}^I, \bar{y}^{\dagger})$, that is, locally, since $\bar{y}^0 = 0$ by (43),

$$f(x^I, 0) + \bar{y}^* \cdot g^*(x^I, 0) \leqq f(\bar{x}^I, 0) + \bar{y}^* \cdot g^*(\bar{x}^I, 0) \tag{151}$$

for all x^I such that $x^{(21)} \geqq 0$.

But $g^*(\bar{x}^I, 0) = 0$ because of (150), and $g^*(x^I, 0)$ in the left member of (151) vanishes for $t^* = \bar{t}^*$. Hence (151) yields, locally and for $t^{(21)} \geqq 0$,

$$f[r^*(\bar{t}^*, t^{**}),\; t^{**}, 0] \leqq f[r^*(\bar{t}^*, \bar{t}^{**}),\; \bar{t}^{**}, 0] \tag{152}$$

which means precisely that $q(\bar{t}^*, t^{**})$ has a local maximum at $\bar{t}^{**}$ subject only to $t^{(21)} \geqq 0$.

Second way: g^0 is empty. In this case (150) is equivalent to

$$g^*(x^I, 0) = 0, \tag{153}$$

$$x^{(21)} \geqq 0. \tag{154}$$

But (153) is necessarily satisfied if $t^* = \bar{t}^*$ and hence can be disregarded. Since $q(t)$

was seen to have a local maximum at $\bar{t}$ subject to (150), it follows that $q(\bar{t}^*, t^{**})$ will have a local maximum at $\bar{t}^{**}$ subject only to $t^{(21)} \geqq 0$.

(b) $$\bar{q}_{t^{**}} = 0 .$$

We have

$$\bar{q}_{t^{**}} = \bar{f}_{x^*}\bar{r}^*_{t^{**}} + \bar{f}_{x^{**}} . \tag{155}$$

We now evaluate the three expressions on the right-hand side of (155). We start with $\bar{r}^*_{t^{**}}$. Noting that

$$f^*\{[r^*(\bar{t}^*, t^{**}),\ t^{**}],\ 0\} = 0 \text{ for all } t^{**} , \tag{156}$$

we obtain by differentiation with respect to t^{**}, using (60) and (69), and evaluating at $t = \bar{t}$,

$$\bar{g}^*_{x^*}\bar{r}^*_{t^{**}} + \bar{g}^*_{t^{**}} = 0 ; \tag{157}$$

in virtue of R_1 this can be solved yielding

$$\bar{r}^*_{t^{**}} = -(\bar{g}^*_{x^*})^{-1}\bar{g}^*_{x^{**}} . \tag{158}$$

To find $\bar{f}_{x^*}$, $\bar{f}_{x^{**}}$, we write the condition that ${}_0\bar{\phi}^I{}_{x^I} = 0$, using equation (41) in the form

$$\begin{aligned} \bar{f}_{x^*} + \bar{g}^*_{x^*}\bar{y}^* &= 0 , \\ \bar{f}_{x^{**}} + \bar{g}^*_{x^{**}}\bar{y}^* &= 0 . \end{aligned} \tag{159}$$

The terms involving g^0 vanish, of course.

Substituting (159) and (158) into (155), we have

$$\bar{q}_{t^{**}} = (-\bar{y}^*\cdot\bar{g}^*_{x^{**}}) + (-\bar{y}^*\cdot\bar{g}^*_{x^*})[-(\bar{g}^*_{x^*})^{-1}\bar{g}^*_{x^{**}}] = (-\bar{y}^*\cdot\bar{g}^*_{x^{**}}) + (\bar{y}^*\cdot\bar{g}^*_{x^{**}}) = 0 . \tag{160}$$

This completes the proof of (71).

REFERENCES

[1] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1951, pp. 481–492.

[2] T. C. Koopmans, "Analysis of production as an efficient combination of activities," edited by T. C. Koopmans, *Activity Analysis of Production and Allocation*, Cowles Commission Monograph No. 13, New York, John Wiley and Sons, 1951, pp. 33–97.

[3] K. J. Arrow and L. Hurwicz, "A gradient method for approximating saddle-points and constrained maxima," P-223, The RAND Corporation, 1951 (hectographed).

[4] G. B. Dantzig, "Maximization of a linear function of variables subject to linear inequalities," *Activity Analysis of Production and Allocation*," edited by T. C. Koopmans, Cowles Commission Monograph No. 13, New York, John Wiley and Sons, 1951, pp. 339–347.

[5] P. A. Samuelson, *Foundations of Economic Analysis*, Cambridge, Harvard University Press, 1948.

[6] A. H. Bowker, "On the norm of a matrix," *Annals of Math. Stat.*, Vol. 18 (1947), pp. 285–288.

[7] F. V. Waugh, "Inversion of the Leontief matrix by power series," *Econometrica*, Vol. 18 (1950), pp. 142–154.

[8] P. A. Samuelson, "Market mechanisms and maximization," P-69, The RAND Corporation, 1949 (hectographed).

[9] L. Hurwicz, "Programming in general spaces," Cowles Commission Discussion Paper, Economics No. 2109, 1954 (hectographed).

[10] E. Hille, *Functional Analysis and Semi-Groups*, New York, American Mathematical Society Colloquium Publications, Vol. 31, 1948.

[11] G. A. Bliss, *Lectures on the Calculus of Variations*, Chicago, The University of Chicago Press, 1946.