COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 80* (Reissue)

(Cowles Commission Discussion Paper: Mathematics 403)
November 7, 1950

LAGRANGE MULTIPLIERS REVISITED

Morton Slater

*CFDP 80, distributed in 1959, was an unchanged reissue of the Cowles Foundation, CCDP: Mathematics 403. It is being reprinted because of its historical interest and continuing demand.

November 1, 1980

LAGRANGE MULTIPLIERS REVISITED

A Contribution to Non-Linear Programming

by Morton Slater

November 7, 1950

1. Introduction

The present paper was inspired by the work of Kuhn and Tucker [1] [1]. These authors transformed a certain class of constrained maximum problems into equivalent saddle value (minimax) problems.

Their work seems to hinge on the consideration of still a third type of problem. A very simple but illustrative form of this problem is the following: let $x \in$ positive orthant of some finite dimensional Euclidean space, and let f and g be real valued functions of x with the property that whenever $f \geq 0$, then also $g \geq 0$; under what conditions can one then conclude that $\exists$ a non-negative constant u such that $uf \leq g$ for <u>all</u> $x \geq 0$?

Kuhn and Tucker showed that if f is concave and differentiable, if g is convex and differentiable, and if the set $\{x: f(x) \geq 0\}$ satisfies certain regularity restrictions, then there does indeed exist such a u.

Two directions for generalization are presented:

First of all, the Kuhn-Tucker argument rests heavily on the

---

1. Numerals in brackets refer to the list of references at the end.

differentiability of the functions, although they express the opinion that their theorems are true without this assumption. Is this the case?

Secondly, the inequality $uf \leqq g$ may be thought of as a relation between f and g. From this point of view, it would appear that a best possible theorem which concludes that $uf \leqq g$ would make assumptions about f only in relation to g, and vice versa.

In this paper it is shown how the second generalization may be partly achieved, and that even with this generalization, differentiability may be dispensed with entirely.

The next section will give a detailed account of the generalized theorem hinted at above. The last section will be devoted to an application of this theorem to transforming a class of constrained minimum problems into equivalent saddle value problems.

2. The Main Theorem

Throughout this section, x, u, and v will be points in the positive orthants of n, k, and h dimensional Euclidean spaces respectively.

<u>Theorem 1</u>. Let $f_1, \ldots, f_k$ and $g_1, \ldots, g_h$, be continuous real valued functions of $x = (x_1, \ldots, x_n)$ for $x \geqq 0$ with the following properties:

1° If at any point x all $f_i(x) \geqq 0$, then for that x some $g_j(x) \geqq 0$

2° $\exists x$ such that all $f_i(x) > 0$

3° If for some x^1 and x^2 $\exists A_i \geqq 0$ $i = 1, \ldots, k$, $B_j \geqq 0$ $j=1, \ldots, h$ and C of arbitrary sign such that

$$\Sigma A_i f_i(x^1) + C = \Sigma B_j g_j(x^1)$$

$$\text{and } \Sigma A_i f_i(x^2) + C = \Sigma B_j g_j(x^2)$$

then for all $x = \theta x^1 + (1-\theta) x^2$ $\quad 0 \leqq \theta \leqq 1$

$$\Sigma A_i f_i(x) + C \geqq \Sigma B_j g_j(x)$$

<u>Assertion</u>: $\exists u_1, \ldots, u_k \geq 0, \; v_1, \ldots, v_h \geq 0 \quad \Sigma v_j = 1$ such that for all $x \geq 0$

$$\Sigma u_i f_i (x) \leqq \Sigma v_j g_j (x)$$

Before proceeding with the sequence of lemmas necessary to prove theorem 1 some discussion of the hypotheses of that theorem is in order.

1^o is clearly essential for the truth of the theorem.

2^o corresponds to the condition of regularity of the constraint set in the Kuhn-Tucker treatment. It may possibly be weakened, but certainly not dispensed with altogether as the example

$$f (x) = -(x - 1)^2 \qquad g (x) = 1 - x$$

shows. Here $f (x)$ is concave, $g (x)$ convex, and 1^o is satisfied. Nevertheless $uf \leqq g$ for some $u \geq 0$ is impossible as is easily verified.

3^o is of course the most controversial hypothesis of all. On the debit side is the fact that any non-negative linear combination of the f_i, say f, must, as a consequence of 3^o, be quasi-concave (i.e. for all real β, $\{x : f (x) \geq \beta\}$ is convex), while any non-negative linear combination of the g_j, say g, must be quasi-convex (i.e. for all real α, $\{x : g (x) \leq \alpha\}$ is convex). This is on the debit side because it puts conditions on the f_i which are independent of the g_j.

Parenthetically we observe that a weaker version of 3^o in which "$A_i \geq 0$" and "$B_j \geq 0$" are replaced by "$A_i > 0$" and "$B_j > 0$" which a priori appears to avoid the above difficulty, actually implies 3^o. (This can easily be shown.)

On the credit side is that if all the f_i are concave and all the g_j convex (the Kuhn-Tucker case), then 3^o is automatically satisfied. What this amounts to is that if a concave function interpolates a convex function at two points, then the concave function dominates the convex function in between.

Moreover, 3^o is satisfied by still other functions. Let f and g be strictly increasing functions of a single real variable x, continuous for $x \geq 0$ and having continuous first and second derivatives for $x > 0$. If in addition $f'(x) > 0$ for all $x > 0$, then by a theorem of M. M. Peixoto [1 , f and g satisfy 3^o if and only if

$$g''(x) \geq \frac{g'(x)}{f'(x)} f''(x) \quad \text{for all} \quad x > 0.$$

Using this theorem, the following examples were easily constructed:

(1) f convex, g convex

$$f(x) = x^2 + x - 2 \qquad g(x) = x^2 - 1 \quad u = \frac{2}{3}$$

(2) f concave, g concave

$$f(x) = \sqrt{x} - 1 \qquad g(x) = \sqrt{x} + x - 2 \quad u = 3$$

(3) Neither f nor g convex or concave

$$f(x) = 2x - \cos x - (2\pi + 1),$$

$$g(x) = x + \sin x - x \cos x - (\pi^2 + \pi)$$

$$u = \pi$$

The main tool in the proof of theorem 1 is the generalized minimax theorem of von Neumann [1] and Kakutani [1], which we shall take as lemma 1.

Lemma 1: Let $\varphi(\zeta,\eta)$ be a continuous real valued function defined for $\xi \in K$ and $\eta \in L$ where K and L are arbitrary bounded closed convex sets of Euclidean spaces R^p and R^q respectively. If for every $\xi^o \in K$ and every real α, the set of all $\eta \in L$ such that $\varphi(\xi^o, \eta) \leq \alpha$ is convex, and if for every $\eta^o \in L$ and every real β, the set of all $\xi \in K$ such that $\varphi(\xi, \eta^o) \geqq \beta$ is convex, then $\exists(\xi^o, \eta^o)$ such that

$$\max_{\xi\in K} \varphi(\xi, \eta^o) = \min_{\eta\in L} \varphi(\xi^o, \eta)$$

Throughout section 2, the functions f_i and g_j will be the functions given in the hypotheses of theorem 1.

Lemma 2: Let $L = \{x : \text{for all } i \;\; f_i(x) \geq 0\}$.

Assertion: $\exists \bar{v} \geq 0$, $\Sigma \bar{v}_j = 1$ such that

$\Sigma \bar{v}_j \, g_j(x) \geq 0$ for all $x \in L$

Proof: Since the $f_i(x)$ are quasi concave, L is closed and convex.

Let $K = \{v : v \geq 0, \;\; \Sigma v_j = 1\}$,

$L_N = \{x : x \in L, \;\; \Sigma x_i \leq N\}$,

and $\varphi(v, x) = \Sigma v_j \, g_j(x)$

By hypothesis 3^o of theorem 1 we may apply lemma 1 to $\varphi(v, x)$ for $v \in K$ and $x \in L_N$. Hence $\exists(v^o, x^o)$ such that

$$\max_{v\in K} \varphi(v, x^o) = \min_{x\in L_N} \varphi(v^o, x) .$$

Using the left hand side of the equality, we see that to prove $\varphi(v^o, x) \geq 0$ for all $x \in L_N$ we need only show that for any $x \in L_N$ $\exists v\in K$ such that $\varphi(v, x) \geq 0$. By hypothesis 1^o of theorem 1 this is clearly the case. Hence $\varphi(v^o, x) \geq 0$ for all $x \in L_N$.

The v^o thus obtained depends on N. Choose $N_n \uparrow \infty$ with n and an associated sequence $\{v^o(n)\}$ such that $\varphi(v^o(n), x) \geq 0$ for all $x \in L_{N_n}$. Since K is compact, some subsequence $\{v^o(n_i)\}$ converges to $\overline{v}$, say. Since $L_{N_{n_i}} \uparrow L$ with i and φ is a continuous function of v we have

$$\varphi(\overline{v}, x) \geq 0 \text{ for all } x \in L \qquad \text{q.e.d.}$$

<u>Notation</u>: The function $\varphi(\overline{v}, x) = \Sigma\, \overline{v}_j\, g(x)$ defined in lemma 2 will be denoted by 'g(x)" for the remainder of section 2.

<u>Lemma 3</u>: Let $f_i(x^o) > 0$ for $i = 1, \ldots, k$ and $g(x^1) < 0$.

<u>Assertion</u>: $\exists i = i_o$ such that $f_{i_o}(x^1) < 0$ and

$$f_{i_o}(x^o)\, g(x^1) - f_{i_o}(x^1)\, g(x^o) \geq 0$$

<u>Proof</u>: We restrict our attention to the line segment joining x^o to x^1; a fortiori 1^o and 3^o are satisfied on the segment. We suppose (for definiteness) that it is oriented thus:

By 1^o, $g(x^o) \geq 0$. Let $\overline{x}$ be the right-most zero of $g(x)$. $(x^o \leq x < x^1)$. By 1^o and the continuity of the $f_i(x)$, $\exists i = i_o$ such that $f_{i_o}(\overline{x}) \leq 0$.

We show first that $f_{i_o}(x^1) < 0$. Suppose false; then $\exists\, \overline{\overline{x}}$ such that $\overline{x} < \overline{\overline{x}} \leq x^1$, $0 \leq f_{i_o}(\overline{\overline{x}}) < f_{i_o}(x^o)$, and $g(\overline{\overline{x}}) < 0$. Hence $\exists A > 0$ and $B < 0$ such that

$$Af_{i_o}(x^o) + B = g(x^o)$$

$$Af_{i_o}(\bar{\bar{x}}) + B = g(\bar{\bar{x}})$$

so that by 3^o, $$Af_{i_o}(x) + B \geqq g(x) \quad \text{for} \quad x^o \leqq x \leqq \bar{\bar{x}}.$$

In particular $$Af_{i_o}(\bar{x}) + B \geqq 0, \quad \text{so that}$$

$$f_{i_o}(x) \geqq -\frac{B}{A} > 0, \quad \text{a contradiction.}$$

Hence $f_{i_o}(x^1) < 0$, and for

$$C = \frac{g(x^o) - g(x^1)}{f_{i_o}(x^o) - f_{i_o}(x^1)} > 0, \qquad D = \frac{f_{i_o}(x^o)g(x^1) - g(x^o)f_{i_o}(x^1)}{f_{i_o}(x^o) - f_{i_o}(x^1)}$$

we have $$C\, f_{i_o}(x^o) + D = g(x^o)$$

$$C\, f_{i_o}(x^1) + D = g(x^1)$$

so that $$C\, f_{i_o}(x) + D \geqq g(x) \qquad \text{for} \quad x^o \leqq x \leqq x^1.$$

In particular

$$C\, f_{i_o}(\bar{x}) + D \geqq 0$$

so that $$D = -\,C\, f_{i_o}(\bar{x}) \geqq 0, \quad \text{and since}$$

$f_{i_o}(x^o) - f_{i_o}(x^1) > 0$, we have

$$f_{i_o}(x^o)g(x^1) - g(x^o)f_{i_o}(x^1) \geqq 0 \qquad \text{q.e.d.}$$

Lemma 4: Let $\varphi(x, u) = u_1 f_1(x) + \dots + u_k f_k(x) - g(x)$

Assertion: $\exists$ a finite positive constant M such that for all $x \geq 0$ $\exists\, u \in L_M = \{u : u \geq 0, \ \Sigma u_i \leq M\}$ such that $\varphi(x, u) \leq 0$.

Proof: If $g(x^1) \geq 0$, $\varphi(x^1, 0) \leq 0$

If $g(x^1) < 0$, proceed as follows: let x^o be the point of the hypothesis 2^o at which all $f_i(x) > 0$. Apply lemma 3 to select i_o such that $f_{i_o}(x^1) < 0$ and

$$f_{i_o}(x^o)g(x^1) - f_{i_o}(x^1)g(x^o) \geqq 0$$

Let
$$u_i = 0 \text{ when } i \neq i_o$$
$$u_{i_o} = \frac{g(x^o)}{f_{i_o}(x^o)}$$

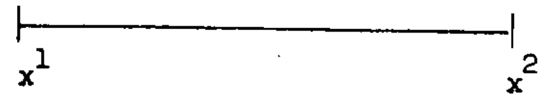
Then
$$\varphi(x\ ,\ u) = \frac{g(x^o)}{f_{i_o}(x^o)}\ f_{i_o}(x^1) - g(x^1)$$
$$= -\frac{f_{i_o}(x^o)g(x^1) - g(x^o)f_{i_o}(x^1)}{f_{i_o}(x^o)} \leq 0$$

Thus if we take $M = \dfrac{g(x^o)}{A}$ where $A = \text{glb } f_i(x^o)$ the lemma is proved. The next lemma will not be needed for the proof of theorem 1, but it is convenient to prove it now.

Lemma 5: Let $f_1(x^1) = \dots = f_\mu(x^1) = 0$, $f_{\mu+1}(x^1), \dots , f_k(x^1) > 0$ and $g(x^1) = 0$.

Assertion: If $f_1(x^2), \dots, f_\mu(x^2) \geq 0$, then $g(x^2) \geq 0$

Proof: As in lemma 3 we consider the functions on the segment joining x^1 to x^2:



If $g(x^2) < 0$, it must have a rightmost zero $\bar{x} < x^2$. We will show that this is impossible.

First of all, by quasi-concavity, $f_i(x) \geq 0$ for all x in the segment and $i \leq \mu$. If any $f_i(x^2) \geqq 0$ for $i > \mu$, then again by quasi-concavity $f_i(x) \geqq 0$ on the whole segment. Finally, suppose some $f_p(x^2) < 0$ for $p > \mu$. Then $\exists A > 0$ such that

$$A\, f_p(x^1) + B = g(x^1)$$

$$A\, f_p(x^2) + B = g(x^2) \quad \text{so that by } 3^\circ$$

$$A\, f_p(x) + B \geqq g(x) \quad \text{for } x^1 \leq x \leq x^2.$$

In particular $\qquad A\, f_p(\bar{x}) + B \geqq 0 \qquad$ so that $f_p(\bar{x}) > 0$.

Hence by continuity, $f_p(x) \geqq 0$ in some right neighborhood of $\bar{x}$.

Combining all this information we see that $\exists\ \bar{\bar{x}} > \bar{x}$ such that all $f\ (\bar{\bar{x}}) \geqq 0$. But by 1°, this implies $g(\bar{\bar{x}}) \geq 0$, a contradiction. q.e.d.

<u>Corollary</u>: If all $f_i(x^1) > 0$ and $g(x^1) = 0$, then $g(x) \geqq 0$ for all x.

The proof of theorem 1 is now easy.

<u>Proof</u>: Choose M as in lemma 4 and $N > 0$ arbitrarily. Consider the function

$$\varphi(x, u) = u_1\, f_1(x) + \dots + u_k\, f_k(x) - g(x)$$

over $\quad x \epsilon \quad K_N = \{x : x \geqq 0 \text{ and } \Sigma x_1 \leq N\}$

$\qquad u \epsilon \quad I_M = \{u : u \geqq 0 \text{ and } \Sigma u_1 \leq M\}$

By lemma 1 and lemma 4 $\exists x^\circ \epsilon K_N, u^\circ \epsilon I_M$ such that

$$\max_{x \epsilon K_N} \varphi(x, u^\circ) = \min_{u \epsilon I_M} \varphi(x^\circ, u) \leq 0$$

Thus $\varphi(x, u^o) \leq 0$ for all $x \in K_N$.

The same kind of compactness argument as in lemma 2 is now used to complete the proof:

Choose $N_n \uparrow \infty$ and an associated sequence $\{u^o(n)\} \subset L_M$ (it is essential to observe that M is independent of N) such that

$$\varphi(x, u^o(n)) \leq 0 \quad \text{for all} \quad x \in K_{N_n}$$

By the compactness of L_M $\exists$ a subsequence $\{u^o(n_i)\}$ converging to $\bar{u}^o$, say. Since $\varphi(x, u)$ is a continuous function of u and $K_{N_{n_i}} \uparrow P = \{x : x \geq 0\}$ we have $\varphi(x, u^o) \leq 0$ for all $x \geq 0$ q.e.d.

3. Applications

<u>Definition</u>: Let $g_1(x), \ldots, g_h(x)$ be any set of functions defined on a set K. A point $x^o \in K$ will be called a <u>minimal point of $g_1, \ldots, g_h$ over K</u> if for all $x \in K$ it is false that for all j $g_j(x) < g_j(x^o)$. A point x^o will be called an <u>essential</u> minimal point of $g_1, \ldots, g_h$ over K if it is minimal and if the deletion of any $g_j(x)$ will cause it to fail to be minimal. A point x^o will be called a <u>strictly</u> minimal point of $g_1, \ldots, g_h$ over K if for all x, if all $g_j(x) \leq g_j(x^o)$ then all $g_j(x) = g_j(x^o)$.

<u>Theorem 2</u>. Let $f_1(x), \ldots, f_k(x)$ and $g_1(x), \ldots, g_h(x)$ be real valued continuous functions which satisfy conditions 2^o and 3^o of theorem 1. Let K be the set $\{x : f_i(x) \geq 0 \;\; i = 1, \ldots, k\}$. Let x^o be an essential minimal point of $g_1(x), \ldots, g_h(x)$ over K.

<u>Assertion</u>: x^o is a strictly minimal point of $g_1(x), \ldots, g_h(x)$ over K.

Proof: The functions $f_i(x)$, $i = 1, \ldots, k$ and $g_j(x) - g(x^o)$ $j = 1, \ldots, h$ satisfy all the hypotheses of theorem 1. Hence by lemma 2 $\exists\ v_1, \ldots, v_h \geq 0$, $\Sigma v_i = 1$ such that

$$\Sigma\, v_j\, g_j(x) \geq \Sigma\, v_j g_j(x^o) \quad \text{for all} \quad x \in K.$$

If for some j, $v_j = 0$, then that $g_j(x)$ may be deleted and x^o will, by the above inequality, remain minimal. Hence all $v_j > 0$. But then, again by the above inequality, if for any $x \in K$ all $g_j(x) \leq g_j(x^o)$, then all $g_j(x) = g_j(x^o)$. q.e.d.

We now proceed to the last theorem, the equivalence of a constrained minimum problem and a certain saddle value problem.

Theorem 3: Let $f_1(x), \ldots, f_k(x)$ and $g_1(x), \ldots, g_h(x)$ be continuous real valued functions of x which satisfy conditions 2^o and 3^o of theorem 1. Let $K = \{x : f_i(x) \geq 0 \ \ i = 1, \ldots, k\}$.

Assertion: x^o is a minimal point for $g_1(x), \ldots, g_h(x)$ over K if and only if $\exists v_1^o, \ldots, v_h^o \geq 0$, $\Sigma v_j^o = 1$, and $u_1^o, \ldots, u_h^o \geq 0$ such that the function

$$\varphi(x, u) = \Sigma v_j^o g_j(x) - \Sigma u_i^o f_i(x) \quad \text{satisfies}$$

$$\varphi(x^o, u) \leq \varphi(x^o, u\ ) \leq \varphi(x, u^o) \tag{1}$$

for all $x \geq 0$. and $u \geq 0$. In other words, $\varphi(x, u)$ has a saddle point at (x^o, u^o).

Proof: Suppose x^o is a minimal point for the $g_j(x)$ over K. Then the functions $f_i(x)$ $i = 1, \ldots, k$ and $g_j(x) - g_j(x_o)$ $j = 1, \ldots, h$ satisfy 1^o, 2^o, 3^o of theorem 1. By lemma 2 we choose $v^o = v_1^o, \ldots, v_h^o)$

so that $\Sigma v_j^o(g_j(x) - g_j(x^o)) \geq 0$ for all $x \in K$. Write $g(x) = \Sigma v_j^o g_j(x)$.

Now suppose $f_1(x^o) = \ldots = f_\mu(x^o) = 0$ and $f_{\mu+1}(x^o), \ldots, f_k(x^o) > 0$. By lemma 5 $g(x) \geq g(x^o)$ for all x such that $f_1(x), \ldots, f_\mu(x) \geq 0$. (In particular, if all $f_i(x^o) > 0$, then $g(x) \geq g(x^o)$ for all x). Hence by theorem 1 $\exists u_1^o, \ldots, u_\mu^o \geq 0$ such that

$$u_1^o f_1(x) + \ldots + u_\mu^o f_\mu(x) \leq g(x) - g(x^o) \quad \text{for all} \quad x \geq 0.$$

Since $f_i(x^o) = 0$ for $i = 1, \ldots, \mu$ this may be rewritten as

$$g(x^o) - \Sigma_1^\mu u_i f_i(x^o) \leq g(x^o) - \Sigma_1^\mu u_i^o f_i(x^o) \leq g(x) - \Sigma_1^\mu u_i^o f_i(x)$$

for all $x \geq 0$ and all $u_i \geq 0$, $i = 1, \ldots, \mu$. Now take $u_{\mu+1}^o = \ldots = u_k^o = 0$. Then

$$g(x^o) - \Sigma_1^k u_i f_i(x^o) \leq g(x^o) - \Sigma_1^k u_i^o f_i(x^o) \leq g(x) - \Sigma_1^k u_i^o f_i(x)$$

for all $x \geq 0$ and all $u = (u\ , \ldots, u\ ) \geq 0$ q.e.d.

Conversely, suppose (1) is satisfied for some v^o and u^o. From the first and last members of the inequality, we find, setting $u = 0$

$$\Sigma_1^h v_j{}^o(g_j(x) - g_j(x^o)) \geq \Sigma_1^k u_i{}^o f_i(x) \quad \text{for all} \quad x \geq 0$$

Hence if $x \in K$, $g_j(x) < g_j(x^o)$ for all j is impossible. This completes the proof.

References

Kakutani, S., 1. "A Generalization of Brouwer's Fixed Point Theorem," Duke Mathematical Journal, Volume 8 (1941), pp 457-459.

Kuhn, H. E., and A. W. Tucker 1. "Non-Linear Programming," RAND Publication LOGS 107 (August 28, 1950).

von Neumann, J., 1. "Uber ein okonomisches Gleichungs-system und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes," Erg. Math. Kolloqu., Volume 8 (1937), pp 73-83.

Peixoto, M. M., 1. "Generalized Convex Functions and Second Order Differential Inequalities," Bulletin of American Mathematical Society, Volume 55 No. 6 (June 1949).