3

THE KUHN-TUCKER THEOREM IN CONCAVE PROGRAMMING

HIROFUMI UZAWA

1. Introduction

In order to solve problems of constrained extrema, it is customary in the calculus to use the method of the Lagrangian multiplier. Let us, for example, consider a problem: maximize $f(x_1, \cdots, x_n)$ subject to the restrictions $g_k(x_1, \cdots, x_n) = 0$ $(k = 1, \cdots, m)$. First, formulate the so-called Lagrangian form

$$\varphi(x, y) = f(x_1, \cdots, x_n) + \sum_{k=1}^{m} y_k g_k(x_1, \cdots, x_n)$$

where unknown $y_1, \cdots, y_m$ are called the Lagrangian multipliers. Then solutions $x_1, \cdots, x_n$ are found among extreme points of $\varphi(x, y)$, with unrestricted x and y, which in turn are characterized as the solutions of

$$\varphi_{x_i}(x, y) = \frac{\partial f}{\partial x_i} + \sum_k y_k \frac{\partial g_k}{\partial x_i} = 0 \qquad (i = 1, \cdots, n) ,$$

$$\varphi_{y_k}(x, y) = g_k(x_1, \cdots, x_n) = 0 \qquad (k = 1, \cdots, m) .$$

This method, although not necessarily true without certain qualifications, has been found to be useful in many particular problems of constrained extrema.

The method is with a suitable modification applied to solve the programming problems also where we are concerned with maximizing a function $f(x_1, \cdots, x_n)$ subject to the restrictions $x_i \geqq 0$ $(i = 1, \cdots, n)$ and $g_k(x_1, \cdots, x_n) \geqq 0$ $(k = 1, \cdots, m)$. Kuhn and Tucker [2] first proved that under some qualifications, concave programming is reduced to finding a saddle-point of the Lagrangian form $\varphi(x, y)$. This Kuhn-Tucker Theorem was further elaborated by Arrow and Hurwicz [1] so that non-concave programming may be handled. In the present chapter we shall, under different qualifications, prove the Kuhn-Tucker Theorem for concave programming.

2. Maximum Problem and Saddle-Point Problem

Let $g(x) = \langle g_1(x), \cdots, g_m(x) \rangle$ (see fn. 1, p. 2) be an m-dimensional vector-valued function and $f(x)$ be a real-valued function, both defined for non-negative vectors $x = \langle x_1, \cdots, x_n \rangle$.

Consider the following

MAXIMUM PROBLEM. *Find a vector* $\bar{x}$ *that maximizes*

(1) $$f(x)$$

G. Giorgi and T.H. Kjeldsen (eds.), *Traces and Emergence of Nonlinear Programming*,
DOI 10.1007/978-3-0348-0439-4_15, © Springer Basel 2014

subject to the restrictions

$$(2) \qquad x \geqq 0 , \quad g(x) \geqq 0 .$$

A vector x will be called *feasible* if it satisfies (2), and a feasible vector $\bar{x}$ maximizing $f(x)$ subject to (2) will be called an *optimum* vector of, or a *solution* to, the problem.

Associated with the Maximum Problem, the *Lagrangian form* is defined by

$$(3) \qquad \varphi(x, y) = f(x) + y \cdot g(x) ,$$

where[1]

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \geqq 0 \text{ and } y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \geqq 0 .$$

A pair of vectors $(\bar{x}, \bar{y})$ is called a *saddle-point* of $\varphi(x, y)$ in $x \geqq 0$, $y \geqq 0$, if

$$(4) \qquad \bar{x} \geqq 0, \quad \bar{y} \geqq 0,$$

$$(5) \qquad \varphi(x, \bar{y}) \leqq \varphi(\bar{x}, \bar{y}) \leqq \varphi(\bar{x}, y) \text{ for all } x \geqq 0 \text{ and } y \geqq 0 ,$$

which may be written as follows:

$$(6) \qquad \varphi(\bar{x}, \bar{y}) = \min_{y \geqq 0} \max_{x \geqq 0} \varphi(x, y) = \max_{x \geqq 0} \min_{y \geqq 0} \varphi(x, y) .$$

SADDLE-POINT PROBLEM. *Find a saddle-point* $(\bar{x}, \bar{y})$ *of* $\varphi(x, y) = f(x) + y \cdot g(x)$.

3. Saddle-Point Implies the Optimality

We are interested in the reduction of a maximum problem to the saddle-point problem of the associated Lagrangian form. First, a proposition will be noted which is true without any qualification on f and g, whenever there exists a saddle-point.

THEOREM 1. *If* $(\bar{x}, \bar{y})$ *is a saddle-point of* $\varphi(x, y)$ *in* $x \geqq 0$, $y \geqq 0$, *then* $\bar{x}$ *is an optimum vector of the maximum problem.*

PROOF. Substituting (3) into (5), we have

$$(7) \qquad f(x) + \bar{y} \cdot g(x) \leqq f(\bar{x}) + \bar{y} \cdot g(\bar{x}) \leqq f(\bar{x}) + y \cdot g(\bar{x}) \quad \text{for all } x \geqq 0,\ y \geqq 0 .$$

---

[1] For two vectors $x = \langle x_1, \cdots, x_n \rangle$ and $u = \langle u_1, \cdots, u_n \rangle$, we shall, as usual, define

$$x \geqq u \text{ if } x_i \geqq u_i , \qquad (i = 1, \cdots, n),$$
$$x \geq u \text{ if } x \geqq u \text{ and } x \neq u ,$$
$$x > u \text{ if } x_i > u_i , \qquad (i = 1, \cdots, n),$$

and $x \cdot u$ stands for the inner product

$$x \cdot u = \sum_{i=1}^{n} x_i u_i .$$

Since the right-hand inequality holds for any $y \geqq 0$, it follows that $g(\bar{x})$ cannot have a negative component, and $\bar{y} \cdot g(\bar{x})$ must be zero:

$$g(\bar{x}) \geqq 0, \qquad \bar{y} \cdot g(\bar{x}) = 0 .$$

Thus the left-hand inequality of (7) may be written as

$$(8) \qquad f(x) + \bar{y} \cdot g(x) \leqq f(\bar{x}) \qquad \text{for all } x \geqq 0 .$$

Since, for any feasible vector x we have $\bar{y} \cdot g(x) \geqq 0$, it follows that $f(x) \leqq f(x) + \bar{y} \cdot g(x) \leqq f(\bar{x})$, which shows that $\bar{x}$ is optimum, q.e.d.

4. The Kuhn-Tucker Theorem

Now a question naturally arises whether, given an optimum vector $\bar{x}$, it is possible to find a vector $\bar{y}$ for which $(\bar{x}, \bar{y})$ is a saddle-point of $\varphi(x, y)$. This, of course, is not true in general, e.g., for convex programming (i.e., where the maximand is a convex function). The following simple example shows that it does not hold even for concave programming:

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

Regarding concave programming, however, the reduction is shown to be possible provided f and g satisfy certain regularity conditions, e.g., the Kuhn-Tucker Constraint Qualification.[2] We shall give sufficient conditions which make the reduction possible.

THEOREM 2. *Suppose that $f(x)$ and $g(x)$ are concave functions on $x \geqq 0$, and $g(x)$ satisfies the following condition (due to M. Slater* [3]) :[3]

(9) *There exists an $x^0 \geqq 0$ such that $g(x^0) > 0$.*

Then a vector $\bar{x}$ is optimum if, and only if, there is a vector $\bar{y} \geqq 0$ such that $(\bar{x}, \bar{y})$ is a saddle-point of $\varphi(x, y)$.

PROOF. Let $\bar{x}$ be optimum. We shall, in the $(m + 1)$-dimensional vector space, define A and B by

$$A = \left\{ \begin{pmatrix} z_0 \\ \\ z \end{pmatrix} ; \begin{pmatrix} z_0 \\ \\ z \end{pmatrix} \leqq \begin{pmatrix} f(x) \\ \\ g(x) \end{pmatrix} \quad \text{for some } x \geqq 0 \right\} ,$$

$$B = \left\{ \begin{pmatrix} z_0 \\ \\ z \end{pmatrix} ; \begin{pmatrix} z_0 \\ \\ z \end{pmatrix} > \begin{pmatrix} f(\bar{x}) \\ \\ 0 \end{pmatrix} \right\} .$$

Since $f(x)$ and $g(x)$ are concave, the set A is convex. Since $\bar{x}$ is opti-

[2] See Kuhn and Tucker [2], p. 483.

[3] A seemingly weaker condition: (9′) For any $u > 0$, there exists a vector $x \geqq 0$ such that $u \cdot g(x) > 0$ is due to S. Karlin. The condition, however, is equivalent to the Slater's condition (9). For the proof, see Chapter 5, pp. 109-10 of the present volume.

mum, A and B have no vector in common. Therefore, by the lemma on the separation of convex sets, there is a non-zero vector $\langle v_0, v\rangle \neq 0$ such that

$$(10) \qquad v_0 z_0 + v \cdot z \leqq v_0 u_0 + v \cdot u \qquad \text{for all } \begin{pmatrix} z_0 \\ z \end{pmatrix} \in A, \begin{pmatrix} u_0 \\ u \end{pmatrix} \in B .$$

By the definition of B, (10) implies $\langle v_0, v\rangle \geqq 0$. Since $\langle f(\bar{x}), 0\rangle$ is on the boundary of B, we also have, by the definition of A,

$$(11) \qquad v_0 f(x) + v \cdot g(x) \leqq v_0 f(\bar{x}) \qquad \text{for all } x \geqq 0 .$$

We have $v_0 > 0$. Otherwise, we have $v \geq 0$ and $v \cdot g(x) \leqq 0$ for all $x \geqq 0$, which contradicts (9).

Let $\bar{y} = v/v_0$. Then

$$(12) \qquad \bar{y} \geqq 0 ,$$

$$(13) \qquad f(x) + \bar{y} \cdot g(x) \leqq f(\bar{x}) \qquad \text{for all } x \geqq 0 .$$

Putting $x = \bar{x}$ in (13), we have $\bar{y} \cdot g(\bar{x}) \leqq 0$. On the other hand, we have

$$(14) \qquad g(\bar{x}) \geqq 0 .$$

Hence

$$(15) \qquad \bar{y} \cdot g(\bar{x}) = 0 .$$

Relations (13), (14), and (15) show that $(\bar{x}, \bar{y})$ is a saddle-point of $\varphi(x, y)$ in $x \geqq 0, y \geqq 0$, q.e.d.

5. A Modification of the Kuhn-Tucker Theorem

Slater's condition (9), however, excludes the case where part of the second half of restriction (2) is

$$h(x) \geqq 0 \qquad \text{and} \qquad -h(x) \geqq 0 ,$$

for linear $h(x)$. In order to make the reduction possible for such cases, we have to modify the Kuhn-Tucker Theorem.

Let sub-sets I and II of $\{1, \cdots, m\}$ be defined by

$$(16) \qquad \mathrm{I} = \{k \,;\ g_k(x) = 0 \qquad \text{for all feasible } x\} ,$$

and

$$(17) \qquad \mathrm{II} = \{1, \cdots, m\} - \mathrm{I} .$$

We shall assume that

(18) *$g_k(x)$ is linear in x, for $k \in \mathrm{I}$.*

(19) *For any i, there is a feasible vector x^i such that*

$$x_i^i > 0 .$$

Then we have as a modification of Theorem 2 the following:

THEOREM 3. *Suppose that* $f(x)$, $g(x)$ *are concave, and* $g(x)$ *satisfies* (18) *and* (19). *Then a vector* $\bar{x}$ *is optimum if, and only if, there is a vector* $\bar{y}$ *such that* $(\bar{x}, \bar{y})$ *is a saddle-point of* $\varphi(x, y)$ *in* $x \geqq 0$ *and* $y \geqq 0$ (II).[4]

PROOF. It is obvious that, if $(\bar{x}, \bar{y})$ is a saddle-point of $\varphi(x, y)$ in $x \geqq 0$ and $y \geqq 0$ (II), then $\bar{x}$ is optimum.

In order to prove the converse we may assume that

$$\frac{dg_k}{dx}, \ k \in I, \ \text{are linear independent.} \tag{20}$$

Let $\bar{x}$ be optimum. We consider two sets A and B defined by

$$A = \left\{ \begin{pmatrix} z_0 \\ z \\ u \end{pmatrix}; \quad z_0 \leqq f(x), \ z_{\mathrm{I}} = g_{\mathrm{I}}(x), \ z_{\mathrm{II}} \leqq g_{\mathrm{II}}(x), \ u \leqq x, \quad \text{for some } x \right\},$$

$$B = \left\{ \begin{pmatrix} z_0 \\ z \\ 0 \end{pmatrix}; \quad z_0 > f(\bar{x}), \ z = 0 \ (\mathrm{I}), \ z > 0 \ (\mathrm{II}) \right\}.$$

Then A and B are convex, and have no point in common. Therefore, there is a vector $\langle v_0, v, w \rangle \neq 0$ such that

$$v_0 \geqq 0, \qquad v \geqq 0 \ (\mathrm{II}), \qquad w \geqq 0, \tag{21}$$

and

$$v_0 f(x) + v \cdot g(x) + w \cdot x \leqq v_0 f(\bar{x}) \qquad \text{for all } x. \tag{22}$$

It now suffices to prove that $v_0 > 0$. If we had assumed that $v_0 = 0$, then

$$v \cdot g(x) + w \cdot x \leqq 0 \qquad \text{for all } x. \tag{23}$$

For any $k \in \mathrm{II}$, there is a feasible vector x^k such that $g_k(x^k) > 0$. Hence

$$v_k = 0 \qquad \text{for } k \in \mathrm{II}. \tag{24}$$

By (19) and (23),

$$w = 0. \tag{25}$$

Using (24) and (25), the inequality (23) may be written as follows:

$$v_{\mathrm{I}} \cdot g_{\mathrm{I}}(x) \leqq 0 \qquad \text{for all } x, \qquad \text{where } v_{\mathrm{I}} \neq 0. \tag{26}$$

Since we have assumed that $g_{\mathrm{I}}(x)$ is linear, (26) implies

[4] The notation $y \geqq 0$ (II) means $y_k \geqq 0$ for all $k \in \mathrm{II}$. The point $(\bar{x}, \bar{y})$ is said to be a saddle-point of $\varphi(x, y)$ in $x \geqq 0$ and $y \geqq 0$ (II) if $\bar{x} \geqq 0$, $\bar{y} \geqq 0$ (II) and (5) holds for any $x \geqq 0$ and $y \geqq 0$ (II).

$$v_{\rm I} \cdot \frac{\partial g_{\rm I}}{\partial x} = 0, \qquad \text{where } v_{\rm I} \neq 0\ , \tag{27}$$

which contradicts (20), q.e.d.

REFERENCES

[1] K. J. Arrow and L. Hurwicz, "Reduction of Constrained Maxima to Saddle-Point Problems," in J. Neyman (ed.), *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley and Los Angeles: University of California Press, 1956, Vol. V, pp. 1-20.

[2] H. W. Kuhn and A. W. Tucker, "Nonlinear Programming," in J. Neyman (ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley and Los Angeles: University of California Press, 1951, pp. 481-92.

[3] M. Slater, "Lagrange Multipliers Revisited: A Contribution to Non-Linear Programming," Cowles Commission Discussion Paper, Math. 403, November 1950.