MEMORANDUM
RM-3858-PR
OCTOBER 1963

A THEOREM OF FRITZ JOHN IN MATHEMATICAL PROGRAMMING

Richard W. Cottle

This research is sponsored by the United States Air Force under Project RAND—contract No. AF 49(638)-700 monitored by the Directorate of Development Planning, Deputy Chief of Staff, Research and Development, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force.

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

DOI 10.1007/978-3-0348-0439-4_6, © Springer Basel 2014

PREFACE

This Memorandum contributes to an aspect of the research program of The RAND Corporation consisting of basic supporting studies in mathematics. Considered is the application of a theorem of Fritz John to mathematical programming.

The author, a consultant to The RAND Corporation, is a Research Assistant at the Operations Research Center, Richmond, California, where some of the earlier research was conducted, partially supported by the Office of Naval Research, Contract No. Nonr-222(83), with the University of California. Preparation of the Memorandum was undertaken under U. S. Air Force Project RAND.

SUMMARY

In this Memorandum, the author specializes a theorem of Fritz John to the case of mathematical programming. It is shown that when a certain multiplier is positive, the well-known Kuhn-Tucker conditions obtain. A sufficient condition for the positivity of this multiplier is proposed.

A THEOREM OF FRITZ JOHN IN MATHEMATICAL PROGRAMMING

1. INTRODUCTION

A 1948 article [1] by F. John, entitled "Extremum Problems with Inequalities as Subsidiary Conditions," appears to be the first paper in which the classical theory of equality-constrained extremization is extended to deal with inequality-constrained extremization. John establishes a theorem on necessary conditions for a minimum and another theorem on sufficient conditions for a relative minimum. The remainder of the paper is devoted to applications of the results. Only the first of these two theorems will be mentioned here.

A more widely known paper is "Nonlinear Programming" by H. W. Kuhn and A. W. Tucker [2] in which John's article is referred to but not discussed in detail. The Kuhn-Tucker paper treats necessary and sufficient conditions for an inequality-constrained maximum.

The purpose of this Memorandum is to point out how the addition of a suitable regularity condition in John's theorem enables one to deduce the "Kuhn-Tucker conditions."

2. THE MATHEMATICAL PROGRAMMING PROBLEM

A typical formulation of the mathematical programming problem is

$$\text{Maximize } f(x) \text{ subject to } x \in E_n , \; g(x) \geq 0 . \qquad (2.1)$$

Further assumptions on the function f and the mapping g yield special types of programming, such as linear, quadratic, concave, etc.

The classical problem referred to earlier is

$$\text{Maximize } f(x) \text{ subject to } x \in E_n , \; g(x) = 0 \qquad (2.2)$$

where f is a differentiable function and

$$g(x) = [g_1(x), \ldots, g_m(x)]^T$$

is a vector-valued differentiable mapping on E_n $(m < n)$. In the method of Lagrange (or undetermined) multipliers, one forms the Lagrangian function

$$L(x,u) = f(x) + u^T g(x) \qquad (2.3)$$

If at a maximum, x^0, the Jacobian matrix $(\partial g_i / \partial x_j)$ has rank m, the following conditions must hold:

$$L_x(x^0,u^0) = 0 \quad \text{some } u^0 \qquad (2.4)$$

$$g(x^0) = 0 . \qquad (2.5)$$

One seeks the solutions of the problem (2.2) among the extrema of the unconstrained Lagrangian $L(x,u)$.*
The conditions (2.4) and (2.5) are necessary, though not sufficient, for an extremum.

We shall state John's theorem below, but with certain notational changes and maximization replacing minimization. The intention is to maintain consistency in problem statements.

Let R be a set of points in E_n and f a real-valued function on R . Let S be a compact metric space. Let $g(x,\sigma) = g_\sigma(x)$ be a real-valued function on $R \times S$. Now define the set

$$R' = \{x \epsilon R \mid g_\sigma(x) \geq 0 \quad \text{all} \quad \sigma \epsilon S\} . \tag{2.6}$$

We seek $x^0 \epsilon R'$ such that

$$f(x^0) = \max_{x \epsilon R'} f(x) . \tag{2.7}$$

Assume that f and $\partial f/\partial x_j$ are continuous on R $(j = 1,\ldots,n)$ and that g and $\partial g/x_j$ are continuous on $R \times S$ $(j = 1,\ldots,n)$. Notice that $R \times S$ can be given a metric space structure.†

---

*See Ref. 3.

†See, for example, Ref. 4, p. 91.

With these notations we are prepared to state

Theorem A:* Let $x^0 \epsilon R'$ be an interior point of R, and let

$$f(x^0) = \max_{x \epsilon R'} f(x) .$$

Then there exists a finite set of points, $\sigma_1, \ldots, \sigma_s \epsilon S$, and numbers, $u_0, u_1, \ldots, u_s$, not all zero, such that

$$g_{\sigma_r}(x^0) = 0 \quad r = 1, \ldots, s \tag{2.8a}$$

$$u_0 \geqq 0 , \; u_1 > 0, \ldots, u_s > 0 \tag{2.8b}$$

$$0 \leqq s \leqq n \tag{2.8c}$$

the function (2.8d)

$$\Phi(x) = u_0 f(x) + \sum_{r=1}^{s} u_r g_{\sigma_r}(x)$$

has a critical point at x^0; i.e., $\Phi_x(x^0) = 0$.

It is important to notice that the "multiplier" u_0 could be zero and that there are no regularity conditions imposed on the constraint set R'.

We shall specialize John's theorem to the case where S is the set $\{1, \ldots, m\}$--which is trivially a compact metric space--and the variables x_j are non-negative.

---

*John, Ref. 1, p. 188.

<u>Theorem B</u>: Let x_0 maximize $f(x)$ constrained by $g_i(x) \geqq 0$, $i = 1,\ldots,m$, and $x_j \geqq 0$, $j = 1,\ldots,n$. Then there exists a semi-positive (i.e., non-negative and non-zero) vector $(u_0,u_1,\ldots,u_m,v_1,\ldots,v_n)^T$ such that

$$u_i g_i(x^0) = 0 \quad i = 1,\ldots,m \tag{2.9a}$$

$$v_j x_j^0 = 0 \quad j = 1,\ldots,n \tag{2.9b}$$

the function (2.9c)

$$\Phi(x) = u_0 f(x) + \sum_{i=1}^{m} u_i g_i(x) + \sum_{j=1}^{n} v_j x_j$$

has a critical point at x^0 .

In order to state the analogous theorem of Kuhn and Tucker, we must first recall the <u>constraint qualification</u>.*

Let x^0 belong to the boundary of the constraint set

$$R' = \{x \epsilon E_n | g_i(x) \geqq 0 \quad i = 1,\ldots,m \ , \ x_j \geqq 0 \ , \ j = 1,\ldots,n\}$$

Let $g^{[1]}(x)$ be the mapping defined by all those component functions of g which vanish at x^0 . Let I_1 consist of those rows of the $n x n$ identity matrix corresponding to components of x^0 which are zero. The Kuhn-Tucker constraint qualification is satisfied at x^0

---

*Ref. 2, p. 483.

if every vector differential dx satisfying the homogeneous linear inequalities

$$g_x^{[1]}(x^0)dx \geqq 0 \ , \ I_1 dx \geqq 0 \qquad (2.10)$$

is tangent to an arc contained in the set R'. This means that to any dx satisfying (2.10) there corresponds a differentiable arc $x = \alpha(\theta)$, $0 \leq \theta \leq 1$, contained in R' such that $x^0 = \alpha(0)$ and some positive scalar λ such that $\alpha'(0) = \lambda dx$.

Theorem C:* Let R' satisfy the constraint qualification. In order that x^0 maximize $f(x)$ subject to $x \varepsilon R'$, it is necessary that x^0 and some $u = (u_1, \ldots, u_m)^T$ satisfy the following conditions:

$$f_x(x^0) + [g_x(x^0)]^T u \leqq 0 \qquad (2.11a)$$

$$(x^0)^T \{f_x(x^0) + [g_x(x^0)]^T u\} = 0 \qquad (2.11b)$$

$$x^0 \geqq 0 \qquad (2.11c)$$

$$g(x^0) \geqq 0 \qquad (2.11d)$$

$$u^T g(x^0) = 0 \qquad (2.11e)$$

$$u \geqq 0 \ . \qquad (2.11f)$$

These relations have also been called the quasi-saddle point conditions [5]. They are necessary conditions of optimality in the program

$$\text{maximize} \quad f(x) \quad \text{subject to} \quad x \varepsilon R' \qquad (2.12)$$

---

*Kuhn-Tucker conditions, Ref. 2, p. 484.

when R' satisfies the constraint qualification. The program (2.12) is called the <u>maximum problem</u> [2].

<u>Theorem 1</u>: Let x^0 solve the problem in Theorem B. If the multiplier u_0 is positive, then the Kuhn-Tucker conditions hold.*

<u>Proof</u>. If $u_0 > 0$ we may assume $u_0 = 1$. Let $u = (u_1,\ldots,u_m)^T$ and $v = (v_1,\ldots,v_n)^T$. Then

$$f_x(x^0) + [g_x(x^0)]^T u = -v \leq 0 \tag{2.13}$$

which is (2.11a). From (2.13) and (2.9b) we get

$$(x^0)^T\{f_x(x^0) + [g_x(x^0)]^T u + v\}$$
$$= (x_0)^T\{f_x(x^0) + [g_x(x^0)]^T u\} = 0$$

which is (2.11b). The remainder of the conditions (2.11) are even more obvious.

<u>3. A SUFFICIENT CONDITION FOR POSITIVE u_0</u>

With all notations as above, let x^0 solve the problem of Theorem B, and let $(u_0,u_1,\ldots,u_m,v_1,\ldots,v_n)^T$ be the associated semi-positive vector of multipliers. $g^{[1]}$ is the mapping composed of components of g which vanish at x^0 . Let $\bar{x}$ be the vector of components x_j of x such that $x_j^0 > 0$.

*A similar result may be found in Ref. 6, p. 227, the English translation of Ref. 4; see also Ref. 2, p. 489.

The <u>regularity condition</u> we shall impose is that the equation

$$y^T[g_{\overline{x}}^{[1]}(x^0)] = 0 \tag{2.14}$$

have no semipositive solution. This condition is slightly more general than the <u>nondegeneracy condition</u> of Ref. 5, which is that $g_{\overline{x}}^{[1]}(x^0)$ be of full rank.

<u>Theorem 2</u>: If g satisfies the regularity condition (2.14), the multiplier u_0 in Theorem B is positive.

<u>Proof</u>. With $\overline{x}$ "evaluated" at x^0, we get $(\overline{x})^0 > 0$, and consequently $\overline{v} = 0$; that is, the corresponding vector of multipliers is zero. Suppose $u_0 = 0$. Then

$$[g_x(x^0)]^T u + v = 0 \tag{2.15}$$

and in particular

$$[g_{\overline{x}}(x^0)]^T u + \overline{v} = [g_{\overline{x}}(x^0)]^T u = 0 . \tag{2.16}$$

Let $u^{[1]}$ be the vector of multipliers corresponding to $g^{[1]}$. Any components in u but not in $u^{[1]}$ must be zero. Therefore, we conclude

$$(u^{[1]})^T[g_{\overline{x}}^{[1]}(x^0)] = 0 . \tag{2.17}$$

Now $u^{[1]}$ is non-negative and cannot be zero, for otherwise u in (2.15) is zero and then so is v.

But this contradicts the semi-positivity of $(u_0, u_1, \ldots, u_m, v_1, \ldots, v_n)^T$. Hence, $u^{[1]}$ is semi-positive. However, (2.17) contradicts our regularity assumption. Therefore, $u_0 > 0$.

REFERENCES

1. John, F., "Extremum Problems with Inequalities as Subsidiary Conditions," Studies and Essays, Courant Anniversary Volume, Interscience, New York, 1948.

2. Kuhn, H. W., and A. W. Tucker, "Nonlinear Programming," Second Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman, ed., University of California Press, Berkeley, 1951, pp. 481-492.

3. Courant, R., Differential and Integral Calculus, Interscience, New York, 1936.

4. Berge, C., Espaces Topologiques, Dunod, Paris, 1959.

5. Arrow, K. J., L. Hurwicz, and H. Uzawa, "Constraint Qualifications in Maximization Problems," Naval Res. Logist. Quart., Vol. 8, No. 2, June 1961, pp. 175-191.

6. Berge, C., Topological Spaces, trans. by E. M. Patterson, MacMillan, New York, 1963.