Appendix

The Lagrange Multiplier Theorem for Max-Min with several constraints

This appendix is based on a paper by J. BRAM [1].

The object of this appendix is to extend the Lagrange-multiplier result of Chapter III (Theorem VII), which was given for the simple constraint $x_1 + \cdots + x_n = X$, $x_i \geqq 0$, $i = 1, \ldots, n$, to the general case of several possibly nonlinear constraints of the form

$$g_j(x) \leqq 0, \quad j = 1, \ldots, m. \tag{1}$$

The result, stated as a theorem below, generalizes also the result of KUHN and TUCKER [6].

We shall maintain the hypotheses of Chapter III that the function $F(x, y)$ and the partial derivatives $F_{x_i}(x, y)$ are continuous in the set of pairs (x, y) for $y \in Y$ and for x in a fixed subset A of Euclidean space containing the set (1). In addition we assume that the $g_j(x)$ are continuously differentiable functions of x in the same set A.

Let x^0 be the point at which $\varphi(x)$ is maximized subject to the conditions (1). We use the usual notation for the gradient:

$$\nabla g_j(x) = \left(\frac{\partial g_j}{\partial x_1}, \ldots, \frac{\partial g_j}{\partial x_n}\right).$$

We denote by J the set of j for which $g_j(x^0) = 0$.

The direction γ is said to be *possible* at x^0 if there is an arc with that direction issuing from x^0 with direction γ in the sense of the Remark following Theorem I of Chapter III, lying entirely in A. The direction γ is *admissible* at x^0 if it is possible and there is an arc isuing from x^0 lying entirely in the constraint space defined by (1).

We denote by Γ the set of directions admissible at x^0. The directional derivative at x^0 exists along these arcs, and does not depend on the particular choice of arc having a direction $\gamma \in \Gamma$, so that

$$D_\gamma \varphi(x) \leqq 0 \quad \text{for} \quad \gamma \in \Gamma. \tag{2}$$

Since these arcs lie in the space (1), then also, if $\gamma \in \Gamma$

$$\nabla g_j(x^0) \cdot \gamma \leqq 0 \quad \text{for} \quad j \in J. \tag{3}$$

We denote the set of all possible γ satisfying (3) by Δ. We have just observed that $\Gamma \subset \Delta$.

We now impose the Kuhn-Tucker constraint qualification, i.e. we assume that $\Gamma = \Delta$. This means that to every possible direction γ satisfying (3) there corresponds an arc issuing from x^0 in that direction and lying in the constrained set (1). This is obviously true in the case when the g_j and A are linear. What it does in the general case is to eliminate pathologies such as certain types of cusps.

Thus with the Kuhn-Tucker constraint qualification assumed we may restate (2) as follows:

$$D_\gamma \varphi(x) \leqq 0 \quad \text{for all} \quad \gamma \in \Delta . \tag{4}$$

Let now W be the set of vectors

$$(F_{x_1}(x^0, y), \ldots, F_{x_n}(x^0, y))$$

for $y \in Y(x^0)$. Formula (12) of Chapter III may be rewritten

$$D_\gamma \varphi(x) = \operatorname*{Min}_{w \in W} \gamma \cdot w , \tag{5}$$

the minimum being achieved.

Let now $\hat{W}$ be the convex hull of W, i.e., the set of vectors $\hat{w}$ representable in the form

$$\hat{w} = \sum_{k=1}^{s} \alpha_k w_k$$

with $\Sigma \alpha_k = 1$, $\alpha_k \geqq 0$, $w_k \in W$, $k = 1, \ldots, s$. Obviously

$$D_\gamma \varphi(x) = \operatorname*{Min}_{\hat{w} \in \hat{W}} \gamma \cdot \hat{w} . \tag{6}$$

Now define the set Δ^* dual to Δ by the condition that $w^* \in \Delta^*$ if and only if

$$w^* \cdot \gamma \leqq 0 \quad \text{for all} \quad \gamma \in \Delta . \tag{7}$$

We are now ready to prove the following theorem.

Theorem. *If x^0 maximizes $\varphi(x)$ subject to the constraints* (1), *and if the Kuhn-Tucker constraint qualification is satisfied, then there exist $\lambda_1, \ldots, \lambda_m$, all non-negative, such that*

$$D_\gamma \varphi(x) \leqq \sum_{j=1}^{m} \lambda_j \gamma \cdot \nabla g_j(x^0)$$

for all possible γ. Further, if $g_j(x^0) < 0$, $\lambda_j = 0$.

Proof. We begin by proving that $\Delta^* \cap \hat{W}$ is not empty. If this intersection were empty, there would be a shortest segment of positive length joining the compact set $\hat{W}$ to the closed set Δ^*. Let the equation of the perpendicular bisector Π of this segment be

$$\gamma^0 \cdot z = \sum_{i=1}^{n} \gamma_i^0 z_i = c . \tag{8}$$

Evidently neither Δ^* nor $\hat{W}$ meets Π: we may therefore suppose that $\gamma^0 \cdot z < c$ on Δ^* and $\gamma^0 \cdot z > c$ on $\hat{W}$. Since $0 \in \Delta^*$, $0 < c$.

Let now $z \in \Delta^*$. Suppose $\varepsilon > 0$. Then $\left(\frac{c}{\varepsilon}\right) z \in \Delta^*$, so that $\gamma^0 \cdot \left(\frac{z}{\varepsilon}\right) < c$, i.e., $\gamma^0 \cdot z < \varepsilon$. It follows that for any $z \in \Delta^*$, $\gamma^0 \cdot z \leqq 0$. Now in particular $\nabla g_j(x^0) \in \Delta^*$, by the definition (7), for all $j \in J$. Hence

$$\gamma^0 \cdot \nabla g_j(x^0) \leqq 0 \quad \text{for all} \quad j \in J . \tag{9}$$

On referring to the definition of Δ [(3) and the first sentence following], we see that (9) implies that $\gamma^0 \in \Delta$. Thus, from (2), $D_{\gamma^0} \varphi(x^0) \leqq 0$, so that, using (5), we have a $w \in W$ such that $\gamma^0 \cdot w \leqq 0$. But this contradicts the statement above that $\gamma^0 \cdot z > c > 0$ for all $z \in \hat{W}$. Hence $\Delta^* \cap \hat{W} \neq \Lambda$.

Now let $z \in \Delta^* \cap \hat{W}$. Since $z \in \Delta^*$, (7) holds with $w^* = z$. Since Δ is defined by the finite set of linear inequalities (3), we may apply the theorem of Farkas [2][14]; there exist non-negative λ_j corresponding to $j \in J$ such that $z = \sum_{j \in J} \lambda_j \nabla g_j(x^0)$. By putting $\lambda_j = 0$ for $j \notin J$, we may write this as follows:

$$z = \sum_{j=1}^{m} \lambda_j \nabla g_j(x^0) .$$

Now we are ready to use the fact that $z \in \hat{W}$. This fact combined with (6) yields, for any possible γ,

$$D_\gamma \varphi(x) = \operatorname*{Min}_{\hat{w} \in \hat{W}} \gamma \cdot w \leqq \gamma \cdot z . \tag{10}$$

But (10) is exactly the statement of the theorem.

[14] This theorem is the following: Write $a = (a_1, \ldots, a_n)$, $u = (u_1, \ldots, u_n)$ and $a \cdot u = \Sigma\, a_i u_i$, Given a collection $a^1, \ldots, a^p$, let S be the set of u such that

$$\begin{gathered} a^1 \cdot u \geqq 0 \\ \vdots \\ a^p \cdot u \geqq 0 . \end{gathered}$$

Suppose that $b \cdot w \geqq 0$ for all $u \in S$. Then there exists a set $\lambda_1, \ldots, \lambda_p$ of non-negative numbers such that

$$b = \sum_{k=1}^{p} \lambda_k a^k .$$

Bibliography

[1] BRAM, J.: The Lagrange multiplier theorem for max-min with several constraints, SIAM Journal, vol. 14, pp. 665—667 (1966).

[2] FARKAS: Über die Theorie der einfachen Ungleichungen. J. reine angew. Math. **124**, 1—27 (1901).

[3] GRAVES, L. G.: The theory of functions of real variables. New York: McGraw-Hill (1946).

[4] KAKUTANI, S.: A generalization of Brouwer's fixed point theorem. Duke Math. J. **8**, No. 3, 457—459 (1941).

[5] KOOPMAN, B. O.: The theory of Search III, optimum distribution of searching effort. Operations Res. **5**, 613—626 (1957).

[6] KUHN, H. W., and A. W. TUCKER: Nonlinear programming. Second Berkeley Symposium on Mathematical Statistics and Probability. University of California Press (1951).

[7] HARLAN MILLS: Marginal values of matrix games and linear programs. Ann. Math., Study No. 38. Linear inequalities and related systems. 183—193 (1956).

[8] DANSKIN, JOHN M.: The theory of max-min, with applications. SIAM Journal, vol. 14, pp. 641—664 (1966).