86

R.T. Rockafellar

ALMOST SURE EXISTENCE OF LAGRANGIAN PRICES VECTORS IN NONLINEAR PROGRAMMING

Estratto dal volume: Giovanni Castellani
and Piera Mazzoleni (edited by), «Mathematical
programming and its economic applications»,
Angeli, Milano, 1981

Franco Angeli Editore

ALMOST SURE EXISTENCE OF LAGRANGIAN PRICE VECTORS IN NONLINEAR PROGRAMMING (°)

by *R. T. Rockafellar*

The characterizations of optimality that are of importance in nonlinear programming fall mainly into two categories. On the one hand there are the differential conditions in terms of gradients, Hessian matrices and the like, or more generally "subgradients" of some sort when the constraint and objective functions are not necessarily differentiable. On the other hand, there are saddle point conditions with respect to the Lagrangian function or an augmented Lagrangian function, and these are typically associated with some dual problem. In both cases the conditions involve Lagrange multipliers that in an economic context may be interpreted as "prices". The conditions, at least in their stronger forms, are typically sufficient for optimality, but they are not necessary unless certain assumptions called "constraint qualifications" are satisfied.

Unfortunately, except for problems of convex type, the constraint qualifications are difficult or impossible to verify. At the least they require tests on an explicit numerical candidate for a solution, but the algorithms for determining a solution cannot find any with exactitude.

(°) Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under grant number 77-3204 at the University of Washington, Seattle.

optimality of x are that x should satisfy the constraints of $(P_{u,\,v})$ and there should exist $y \in R_{+}^{m}$ satisfying

$$(1)\qquad \begin{aligned} &\nabla g_0(x) + \Sigma_{i=1}^{m} y_i \nabla g_i(x) = v , \\ &y_i = 0 \text{ for } i \notin I(x,\, u) , \end{aligned}$$

where $I(x, u)$ is the set of active constraint indices (the inequalities that hold as equations). Among the many constraint qualifications that are known to imply the necessity of these conditions, the simplest and most powerful is

(2) the vectors $\nabla g_i(x)$ for $i \in I(x, u)$ are linearly independent .

If the functions are twice continuously differentiable, there are further optimality conditions in terms of the Hessian matrix

$$H(x,\, y) = \nabla^2 g_0(x) + \sum_{i=1}^{m} y_i \nabla^2 g_i(x) .$$

The classical *strong second-order conditions* are that x should satisfy the first-order conditions and (2) plus the properties

(3) $y_i > 0$ for all $i \in I(x, u)$,

(4) $z \cdot H(x, y)z > 0$ for every nonzero $z \in R^n$ satisfying $z \cdot \nabla g_i(x) = 0$ for all $i \in I(x, u)$.

These conditions are sufficient for x to be an isolated locally optimal solution to $(P_{u,\,v})$.

They imply further that small changes in u and v affect x and y in a differentiable manner. Specifically, there exist open neighborhoods U and V of u and v and functions $\xi : U \times V \to R^n$ and $\eta : U \times V \to R^m$ of class C^1 such that $\xi(u, v) = x$, $\eta(u, v) = y$, and for all $(u', v') \in U \times V$ the vectors $x' = (u', v')$ and $y' = \xi(u', v'$ satisfy the strong second-order conditions for $(P_{u',v'})$, so that

x' is locally optimal for $(P_{u',v'})$. This is a property of obvious importance in economic analysis.

The following results may be derived from the well known theorem of Sard (cf. [2]) on the critical values of a differentiable mapping. The proofs are contained in a forthcoming paper of Spingarn and Rockafellar [3].

THEOREM 1. Suppose C *is open in* R^n, g_0 *is of class* C^1, *and* $g_1, \ldots, g_m$ *are of class* C^{n+1}. *Fix any* $v \in R^n$. *Then for all* u *except in a negligeable subset of* R^m, $(P_{u,v})$ *is such that every feasible solution* x *satisfies the constraint qualification* (2), *and hence every locally optimal solution* x *satisfies the constraint qualification* (2), *and hence every locally optimal solution* x *satisfies the first-order conditions for some* $y \in R^m_+$.

THEOREM 2. Suppose C *is open in* R^n *and* $g_0, g_1, \ldots, g_m$ *are of class* C^2. *Then for all* (u,v) *except in a negligeable subset of* $R^m \times R^n$, $(P_{u,v})$ *is such that every locally optimal solution* x *and every* $y \in R^m_+$ *satisfying the first-order conditions with* x, *the strong second-order conditions actually hold.*

A tricky feature of the proofs is the need for handling inequalities and complementary slackness conditions. These do not fit the standard mold for applications of Sard's theorem. Note that the differentiability requirements in Theorem 2 are weaker than those of Theorem 1, as far as the constraints are concerned. Thus one can get away with less differentiability if there is some other property guaranteeing that the first-order conditions are satisfied by all locally optimal solutions.

COROLLARY 1. Suppose C *is open,* g_0 *is of class* C^2, *and* $g_1, \ldots, g_m$ *are of class* C^{n+1}. *Then for all* (u, v) *except in a* *negligeable subset of* $R^m \times R^n$, $(P_{u,v})$ *is such that every locally optimal solution* x *satisfies the strong second-order conditions.*

This conclusion is also valid if $g_1, \ldots, g_m$ *are of class* C^2 *and convex (and* C *is convex).*

The justification of the last assertion of Corollary 1 lies in the fact that when the constraint system is convex, the set of vectors u for which $(P_{u,v})$ has feasible solutions is a convex set in R^m whose nonempty interior consists of those for which $(P_{u,v})$ satisfies the Slater condition. The Slater condition serves as a constraint qualification for convex systems. The boundary of a convex set in R^m is negligeable. Thus again, the set of vectors (u, v) such that $(P_{u,v})$ has a locally optimal solution not satisfying the first-order conditions is negligeable.

It must be emphasized that these results represent only a beginning step in the construction of an adequate theory, although they do illustrate some of the tools available and the difficulties involved. A simple example for which Theorems 1 and 2 do not yield a significant conclusion occurs when x is naturally constrained to be nonnegative. If this requirements is represented by a system of inequalities

$$0 \geq g_{m+j}(x) = -e_j \cdot x \text{ for } j = 1, \ldots, n,$$

the assertions of "typical" behavior will be relative to a family of problems involving additional parameters u_{m+j}. Thus we will be looking at problems that include constraints of the form

$$x_j \geq -u_{m+j} \quad \text{for} \quad j = 1, \ldots, n.$$

But among these, the subfamily with the desired constraints

$$x_j \geq 0 \quad \text{for } j = 1, \ldots, n$$

is negligeable!

An alternative way to treat a constraint like $x \geq 0$ would be to incorporate it into the set C. But then C would not be open, so the very nature of the first and second-order conditions for optimality would be affected. Actually, for a constraint as simple as $x \geq 0$, this would not lead to an impass. The general implication is clear however. One needs a theory of second-order conditions that can be applied to $(P_{u,v})$ in the case of a reasonable class of sets C capable of incorporating smooth but nonlinear constraints that are purely "structural", i.e. not suitably viewed as parameterized.

Such a theory has been developed by Spingarn [1] for sets he calls *cyrtohedra*. It yields much stronger versions of Theorems 1 and 2.

Many questions remain open, though. For instance, when is it true that the set of (u, v) for which $(P_{u,v})$ has a multiplicity of globally optimal solutions is negligeable? An answer to this would tie in with the theory of augmented Lagrangians.

Recall that the augmented Lagrangian function for $(P_{u,v})$ has the form

$$L(x,y,r) = g_0(x) - v\cdot x + \sum_{i=1}^{m} \theta(g_i(x)-u_i, y_i, r)$$

where $r > 0$, $y \in R^m$ (*not* necessarily R^m_+) and

$$\theta(t_i, y_i, r) = \begin{cases} y_i t_i + r t_i^2 & \text{if} \quad t_i \geq -y_i/2r \\ -y_i^2/4r & \text{if} \quad t_i \leq -y_i/2r \end{cases}$$

It has been shown by Rockafellar [4, Theorem 6] that for any globally optimal solution x that satisfies the second-order conditions (even the "weak form") and is "strongly *unique*", there exist y and r such that (x, y, r) is a global saddle point of L. While the global property remains unsettled in the context of Theorems 1 and 2, one can at least draw the following conclusion.

COROLLARY 2. Suppose C *is open,* g_0 *is of class* C^2 *, and* $g_1, \ldots, g_m$ *are of class* C^{n+1}*. Then for all* (u, v) *except in a negligeable subset of* $R^m \times R^n$, $(P_{u,v})$ *is such that every locally optimal solution* x *corresponds to a local saddle point* (x, y, r) *of the augmented Lagrangian.*

This conclusion is also valid if $g_1, \ldots, g_m$ *are of class* C^2 *and convex (and* C *is convex).*

The converse assertion, namely that x is locally optimal if (x, y, r) is a local saddle point of the augmented Lagrangian, is true [4] without any assumption on C, $g_0, \ldots, g_m$.

REFERENCES

[1] J.E. Spingarn, *Generic Conditions for Optimality in Nonlinear Programming Problems*, Thesis, University of Washington, 1977. (Author's address: Department of Mathematics, Georgia Tech., Atlanta, Georgia).

[2] R. Abraham, J. Robbin, *Transversal Mappings and Flows*, W. H. Benjamin, Inc., 1967.

[3] J. E. Spingarn, R.T. Rockafellar, "The generic nature of optimality conditions in nonlinear programming", *Math. of Operations Research*, to appear.

[4] R. T. Rockafellar, "Augmented Lagrange multiplier functions and duality in nonlinear programming", *SIAM J. Control*, 12, 1974, pp.268-285.