JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 77, No. 1, APRIL 1993

First-Order Conditions for Isolated Locally Optimal Solutions[1]

A. SHAPIRO[2] AND F. AL-KHAYYAL[2]

Communicated by O. L. Mangasarian

Abstract. There are well-known first-order sufficient conditions for a point x_0 to be a strict locally optimal solution of a nonlinear programming problem. In this paper, we show that these conditions also guarantee that x_0 is an isolated stationary point of the considered program provided a constraint qualification holds. This result has an interesting application to finite convergence of algorithms along the lines suggested by Al-Khayyal and Kyparisis.

Key Words. First-order optimality conditions, tangent cones, monotone multifunctions, finite convergence of algorithms.

1. Introduction

Consider the nonlinear programming problem

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & x \in S, \end{aligned} \tag{1}$$

where the set S is given by constraints

$$S = \{x \in \mathbb{R}^n : g_i(x) = 0,\ i = 1, \ldots, q;\ g_i(x) \leq 0,\ i = q+1, \ldots, p\}. \tag{2}$$

Suppose that the functions $f(x)$ and $g_i(x)$, $i = 1, \ldots, p$, are continuously differentiable and denote by $I(x_0)$ the index set of active inequality constraints at $x_0 \in S$. Then, it is well known that the condition

$$-\nabla f(x_0) \in \operatorname{int}\left\{ y : y = \sum_{i=1}^{q} \lambda_i \nabla g_i(x_0) + \sum_{i \in I(x_0)} \lambda_i \nabla g_i(x_0);\ \lambda_i \geq 0,\ i \in I(x_0) \right\} \tag{3}$$

[1]This research was supported in part by National Science Foundation Grant DDM-91-14489.

[2]Associate Professor, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

0022-3239/93/0400-0189$07.00/0

is sufficient for the point x_0 to be a strictly optimal local solution of (1) in the sense that

$$f(x) > f(x_0), \qquad \text{for all } x \in S,\ x \neq x_0,$$

in a neighborhood of x_0. However, condition (3) alone does not imply that x_0 is an isolated local solution of (1). Let us consider the following example which is a slight modification of an example due to Robinson (Ref. 1):

$$\min \quad x_1, \tag{4a}$$

$$\text{s.t.} \quad x_1^4 \sin^2(1/x_1) + x_2^2 \leq 0, \tag{4b}$$

$$-x_1 + x_2 \leq 0, \tag{4c}$$

$$-x_1 - x_2 \leq 0. \tag{4d}$$

The point $x_0 = (0, 0)$ satisfies condition (3), but it is not an isolated local solution of (4). This is because the feasible set S of the program (4) is formed by a sequence of locally separated points converging to $(0, 0)$ with every point of this sequence being a locally optimal solution of (4).

In this paper, we show that, indeed, x_0 is an isolated locally optimal solution of (1) if, in addition to condition (3), a constraint qualification holds (for example, the Mangasarian–Fromovitz condition). Somewhat surprisingly, second-order conditions will not be required. This can be compared with the results of Robinson (Ref. 1) where, in the absence of assumption (3), second-order sufficient conditions were employed. We also mention some interesting applications of our results to finite convergence of algorithms along the lines suggested by Al-Khayyal and Kyparisis (Ref. 2).

Although our investigation is motivated by finite-dimensional applications, finite dimensionality of the considered linear space is not essential for derivation of our main results. Therefore, we assume henceforth that S is a closed subset of a Banach space X and denote by X^* its dual space of bounded linear functionals. For $x \in X$ and $y \in X^*$, we use the notation $\langle y, x\rangle$ for the value $y(x)$. We denote the distance from a point x to the set S by $\operatorname{dist}(x, S)$.

It should be mentioned that condition (3) and its abstract analogue given in (6) are quite restrictive. In the finite-dimensional space $\mathbb{R}^n$, condition (3) can hold only if

$$q + \operatorname{card}\{I(x_0)\} \geq n.$$

Condition (6) requires the interior of the normal cone $N_S(x_0)$ to be nonempty. Moreover, in infinite-dimensional spaces, (6) never holds if the set S is defined by a finite number of constraints.

2. Nearly Convex Sets and First-Order Optimality Conditions

Before giving the main results of this section, we need to introduce the following definitions. It will be assumed that the function $f(x)$ is continuously (Fréchet) differentiable with the corresponding derivative $Df(x) \in X^*$. Recall that the contingent (Bouligand) cone $T_S(x)$ to the set S at a point $x \in S$ is formed by vectors y such that there exist sequences $x_n \in S$ and $t_n \in \mathbb{R}_+$, with $x_n \to x$, $t_n \to 0^+$, and $t_n^{-1}(x_n - x) \to y$. It is well known that, if x_0 is a locally optimal solution of the program (1), then

$$\langle Df(x_0), y\rangle \geq 0, \qquad \text{for all } y \in T_S(x_0).$$

These first-order necessary conditions can be formulated in the equivalent form

$$-Df(x_0) \in N_S(x_0), \tag{5}$$

where

$$N_S(x) = \{z \in X^*: \langle z, y\rangle \leq 0, \text{for all } y \in T_S(x)\}$$

denotes the polar cone of the contingent cone $T_S(x)$.

It is also not difficult to show that the condition

$$-Df(x_0) \in \operatorname{int} N_S(x_0) \tag{6}$$

is sufficient for x_0 to be a strictly optimal local solution of (1) provided the cone $T_S(x_0)$ has the following approximating property:

$$\operatorname{dist}(x - x_0, T_S(x_0)) = o(\|x - x_0\|), \qquad x \in S. \tag{7}$$

Note that, in finite-dimensional spaces, condition (7) always holds.

This can be easily proved by standard arguments of compactness. That is, suppose that (7) is false. This will imply the existence of a sequence $x_n \in S$ converging to x_0 and $\epsilon > 0$ such that

$$\|x_n - x_0\|^{-1} \operatorname{dist}(x_n - x_0, T_S(x_0)) \geq \epsilon.$$

Moreover, because of the compactness of the unit ball in $\mathbb{R}^n$, we can assume that

$$y_n = \|x_n - x_0\|^{-1}(x_n - x_0)$$

converges to a vector y. It follows that $y \in T_S(x_0)$. On the other hand,

$$\operatorname{dist}(y_n, T_S(x_0)) \geq \epsilon,$$

and hence $y \notin T_S(x_0)$, a contradiction.

Definition 2.1. We say that the set S is nearly convex at a point $x_0 \in S$ if there exists a function $K(x, y)$ tending to zero as $x \to x_0$, $y \to x_0$, and such that, for all $x, y \in S$,

$$\operatorname{dist}(y - x, T_S(x)) \leq K(x, y)\|y - x\|. \tag{8}$$

Remark 2.1. Condition (8) can be viewed as a uniform version of the approximating property (7). If the set S is convex, then $y-x \in T_S(x)$ for any y and x in S, and hence condition (8) follows.

We can now formulate the first result of our paper.

Theorem 2.1. Suppose that $f(x)$ is continuously differentiable, that S is nearly convex at x_0, and that condition (6) holds. Then, x_0 is an isolated locally optimal solution of the problem (1).

Before giving a proof of Theorem 2.1, we shall need the result in the following lemma.

Lemma 2.1. Suppose that S is nearly convex at x_0. Then, the multifunction $x \to N_S(x)$ is almost monotone near x_0 in the sense that, for all $x_1, x_2 \in S$ and $y_1 \in N_S(x_1)$, $y_2 \in N_S(x_2)$,

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq -[K(x_1, x_2)\|y_1\| + K(x_2, x_1)\|y_2\|]\|x_1 - x_2\|. \tag{9}$$

Proof. Consider $x_1, x_2 \in S$ and $y_1 \in N_S(x_1)$. Since S is nearly convex, we have by the definition that, for any $\epsilon > 0$, there is an $\bar{x}_2 \in x_1 + T_S(x_1)$ such that

$$\|x_2 - \bar{x}_2\| \leq K(x_1, x_2)\|x_1 - x_2\| + \epsilon. \tag{10}$$

Because $\bar{x}_2 - x_1 \in T_S(x_1)$, it follows that

$$\langle y_1, \bar{x}_2 - x_1 \rangle \leq 0,$$

and by (10) we also have

$$\langle y_1, x_2 - \bar{x}_2 \rangle \leq \|y_1\|[K(x_1, x_2)\|x_1 - x_2\| + \epsilon].$$

Since $\epsilon > 0$ was arbitrary, this implies that

$$\langle y_1, x_2 - x_1 \rangle \leq K(x_1, x_2)\|y_1\| \, \|x_1 - x_2\|. \tag{11}$$

Similarly, for $y_2 \in N_S(x_2)$,

$$\langle y_2, x_1 - x_2 \rangle \leq K(x_2, x_1)\|y_2\| \, \|x_1 - x_2\|. \tag{12}$$

Inequality (9) then follows from (11) and (12). □

Proof of Theorem 2.1. By contradiction, suppose that x_0 is not an isolated local solution. We have then that there is a sequence $x_n \in S$ of local solutions converging to x_0. Now, let $\{z_n\}$ be a sequence in the dual space X^* such that $\|z_n\| = 1$ and

$$\langle z_n, x_0 - x_n \rangle = \|x_0 - x_n\|. \tag{13}$$

Since $-Df(x_0)$ belongs to the interior of $N_S(x_0)$, there is an $\epsilon > 0$ such that, if the distance from a point v to $-Df(x_0)$ is less than or equal to ϵ, then $v \in N_S(x_0)$. Consider now

$$v_n = -Df(x_0) - \epsilon z_n.$$

We have that $v_n \in N_S(x_0)$. Also, by the first-order necessary conditions [see (5)],

$$-Df(x_n) \in N_S(x_n).$$

It then follows from result (9) of Lemma 2.1 that

$$\begin{aligned}&\langle v_n + Df(x_n), x_0 - x_n\rangle\\ &\geq -[K(x_0, x_n)\|v_n\| + K(x_n, x_0)\|Df(x_n)\|]\|x_n - x_0\|.\end{aligned}$$

Moreover, since sequences $\{v_n\}$ and $\{Df(x_n)\}$ are bounded, we have

$$\langle v_n + Df(x_n), x_0 - x_n\rangle \geq o(\|x_0 - x_n\|).$$

Then, since $v_n = -Df(x_0) - \epsilon z_n$, we have

$$-\epsilon\langle z_n, x_0 - x_n\rangle \geq \langle Df(x_0) - Df(x_n), x_0 - x_n\rangle + o(\|x_0 - x_n\|);$$

hence, by continuity of $Df(x)$ and because of (13), we obtain

$$-\epsilon\|x_0 - x_n\| \geq o(\|x_0 - x_n\|).$$

But, for n large enough, the last inequality contradicts itself, and hence the proof is complete. □

Remark 2.2. Recall that a point $x_0 \in S$ is a stationary point of the program (1) if condition (5) holds. We have shown that, under the assumptions of Theorem 2.1, x_0 is an isolated stationary point of the program (1).

Let us suppose now that the set S is given in the form

$$S = \{x \in X: g(x) \in K\}, \tag{14}$$

where $g(x)$ is a continuously differentiable mapping from X into a Banach space Y and K is a closed convex cone in Y. A point x_0 is a regular point of $g(x)$, in the sense of Robinson (Ref. 3), if

$$0 \in \text{int}\{g(x_0) + Dg(x_0)X - K\}. \tag{15}$$

For example, if

$$X = \mathbb{R}^n, \qquad Y = \mathbb{R}^p, \qquad -K = \{0\} \times \mathbb{R}_+^{p-q},$$

then the set S defined in (14) is the same as the one given in (2), and the regularity condition (15) is equivalent to the Mangasarian-Fromovitz (Ref. 4) constraint qualification.

Theorem 2.2. Suppose that the mapping $g(x)$ is continuously differentiable and that $x_0 \in S$ is a regular point of $g(x)$. Then, the set S is nearly convex at x_0.

Proof. For a point $x \in S$, consider the cone

$$T_S(x) = \{y: Dg(x)y \in T_K(g(x))\}. \tag{16}$$

It is known that, if x is a regular point of $g(x)$, then this cone is the contingent cone to S at x. This is a simple consequence of the Robinson–Ursescu (Refs. 5, 6) stability theorem and the fact that the tangent cone $T_K(g(x))$ (in the sense of convex analysis, e.g., Ref. 7) to the convex cone K at the point $g(x)$ is also its contingent cone (cf. Ref. 3, pp. 504–505, and Ref. 8, Section 2). Also, it follows from regularity of x_0 that all points x sufficiently close to x_0 are regular (e.g., Ref. 9, Theorem 2.2).

Let us observe now that, since K is convex,

$$K - g(x) \subset T_K(g(x)),$$

and hence $T_S(x)$ contains the set

$$C(x) = \{y: Dg(x)y \in K - g(x)\} = \{y: g(x) + Dg(x)y \in K\}.$$

We have then that

$$\text{dist}(y - x, T_S(x)) \leq \text{dist}(y - x, C(x)).$$

Again by the Robinson–Ursescu stability theorem, there is a neighborhood N of x_0 and a constant α such that, for all $x, y \in S \cap N$,

$$\text{dist}(y - x, C(x)) \leq \alpha\, \text{dist}(g(x) + Dg(x)(y - x), K).$$

Moreover, since $g(y) \in K$, we have

$$\text{dist}(g(x) + Dg(x)(y - x), K) \leq \|g(x) + Dg(x)(y - x) - g(y)\|,$$

and hence,

$$\text{dist}(y - x, T_S(x)) \leq \alpha \|g(y) - g(x) - Dg(x)(y - x)\|.$$

Since $g(x)$ is continuously differentiable, we have by the mean value theorem that

$$\lim_{y, x \to x_0} \|y - x\|^{-1} \|g(y) - g(x) - Dg(x)(y - x)\| = 0.$$

Together with the last inequality, this implies (8), and hence the proof is complete. □

As a consequence of Theorems 2.1 and 2.2, any point $x_0 \in S$ satisfying both condition (3) and the Mangasarian-Fromovitz constraint qualification is an isolated stationary point of the mathematical programming problem (1).

Under the assumption that S is convex, the main result in Ref. 2 (see also Ref. 10) is that a convergent algorithm for finding a local solution of program (1) will terminate after solving finitely many auxiliary problems of the form

$$\begin{aligned} &\min \quad \nabla f(x_k)^T x, \\ &\text{s.t.} \quad x \in S, \end{aligned}$$

where x_k is the sequence of generated points converging to x_0. It follows from our results that the convexity of S can be relaxed and replaced by the assumption of near convexity while maintaining the finite termination property.

Finally, we compare condition (6) with the notion of a sharp minimum (or equivalently, a strongly unique local minimum) discussed in Refs. 10 and 11. It is said that problem (1) has a sharp minimum at x_0 if

$$f(x) \geq f(x_0) + \alpha \|x - x_0\|, \tag{17}$$

for all $x \in S$ near x_0 and some $\alpha > 0$ (cf. Ref. 12). We show now that the above condition of sharp minimum essentially is equivalent to condition (6). That is, we show that conditions (6) and (7) imply that $f(x)$ attains a sharp minimum at x_0; conversely, if $f(x)$ has a sharp minimum at x_0, then condition (6) follows.

Suppose that conditions (6) and (7) hold. We have then that there is an $\epsilon > 0$ such that the ball of radius ϵ and centered at $-Df(x_0)$ is contained in $N_S(x_0)$. It follows that

$$\langle Df(x_0), y\rangle \geq \epsilon \|y\|, \qquad \text{for all } y \in T_S(x_0).$$

Taking the first-order Taylor expansion of $f(x)$ at x_0 and using approximating property (7), we obtain that condition (17) follows (e.g., with $\alpha = \epsilon/2$) for all $x \in S$ near x_0.

Conversely, suppose that $f(x)$ attains a sharp minimum at x_0. Consider $y \in T_S(x_0)$, and let $x_n \in S$ be a sequence converging to x_0 such that $t_n^{-1}(x_n - x_0) \to y$ for some $t_n \to 0^+$. It follows from (17) that, for all n large enough,

$$t_n^{-1}[f(x_n) - f(x_0)] \geq \alpha \|t_n^{-1}(x_n - x_0)\|. \tag{18}$$

By passing to the limit in (18), we obtain

$$\langle Df(x_0), y\rangle \geq \alpha \|y\|, \qquad \text{for all } y \in T_S(x_0).$$

This implies that the ball of radius α and centered at $-Df(x_0)$ is contained in $N_S(x_0)$, and hence condition (6) follows.

References

1. ROBINSON, S. M., *Generalized Equations and Their Solutions, Part 2: Applications to Nonlinear Programming*, Mathematical Programming Study, Vol. 19, pp. 200–221, 1982.
2. AL-KHAYYAL, F. A., and KYPARISIS, J., *Finite Convergence of Algorithms for Nonlinear Programs and Variational Inequalities*, Journal of Optimization Theory and Applications, Vol. 70, pp. 317–330, 1991.
3. ROBINSON, S. M., *Stability Theory for Systems of Inequalities, Part 2: Differentiable Nonlinear Systems*, SIAM Journal on Numerical Analysis, Vol. 13, pp. 497–513, 1976.
4. MANGASARIAN, O. L., and FROMOVITZ, S., *The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints*, Journal of Mathematical Analysis and Applications, Vol. 7, pp. 37–47, 1967.
5. ROBINSON, S. M., *Regularity and Stability for Convex Multivalued Functions*, Mathematics of Operations Research, Vol. 1, pp. 130–143, 1976.
6. URSESCU, C., *Multifunctions with Convex Closed Graph*, Czechoslovak Mathematical Journal, Vol. 25, pp. 438–441, 1975.
7. AUBIN, J. P., and EKELAND, L., *Applied Nonlinear Analysis*, Wiley, New York, New York, 1984.
8. MAURER, H., and ZOWE, J., *First- and Second-Order Necessary and Sufficient Optimality Conditions for Infinite-Dimensional Programming Problems*, Mathematical Programming, Vol. 16, pp. 98–110, 1979.
9. ALT, W., *Stability of Solutions for a Class of Nonlinear Cone Constrained Optimization Problems, Part 1: Basic Theory*, Numerical Functional Analysis and Optimization, Vol. 10, pp. 1053–1064, 1989.
10. FERRIS, M. C., and MANGASARIAN, O. L., *Minimum Principle Sufficiency*, Mathematical Programming, Vol. 57, pp. 1–14, 1992.
11. BURKE, V. J., and FERRIS, M. C., *Weak Sharp Minima in Mathematical Programming*, Technical Report 1050, Computer Sciences Department, University of Wisconsin-Madison, 1991.
12. POLYAK, B. T., *Introduction to Optimization*, Optimization Software, Publications Division, New York, New York, 1987.