J. Austral. Math. Soc. (*Series A*) 41 (1986), 188–192

QUASI-REGULARITY IN OPTIMIZATION

KUNG-FU NG and DAVID YOST

(Received 16 October 1984; revised 25 February 1985)

Communicated by B. Mond

Abstract

The notion of quasi-regularity, defined for optimization problems in $\mathbf{R}^n$, is extended to the Banach space setting. Examples are given to show that our definition of quasi-regularity is more natural than several other possibilities in the general situation. An infinite dimensional version of the Lagrange multiplier rule is established.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*): 49 B 27.

In a classical situation one is concerned to minimize a real-valued function on $\mathbb{R}^n$ subject to

$$(*) \qquad g_i(x) \leqslant 0 \qquad (i = 1, 2, \ldots, m).$$

Let A denote the feasible set (consisting of all x satisfying $(*)$), and let $x_0 \in A$. The *tangent cone* of A at x_0 is denoted by $\mathrm{TC}[A; x_0]$ and is defined to be the cone (i.e. the positively homogeneous set) generated by all unit vectors h for each of which there exists a sequence $\{a_q\}_{q=1}^{\infty}$ in A with $a_q \to x_0$ and $(a_q - x_0)\|a_q - x_0\|^{-1} \to h$ when $q \to \infty$. The (*outer*) *normal cone* is denoted by $\mathrm{NC}[A; x_0]$ and is defined to be the polar of $\mathrm{TC}[A; x_0]$:

$$\mathrm{NC}[A; x_0] = \{\xi \in \mathbb{R}^n : \xi \cdot h \leqslant 0 \ \forall h \in \mathrm{TC}[A; x_0]\}.$$

Thus, if i is active at x_0 in the sense that $g_i(x_0) = 0$, then the gradient $\nabla g_i(x_0)$ is an outer normal because

$$\nabla g_i(x_0) h = \lim_{q \to \infty} \frac{g_i(a_q) - g_i(x_0)}{\|a_q - x_0\|} \leqslant 0.$$

This shows that if

$$S_0 = \left\{ \sum_{i=1}^{m} \lambda_i \nabla g_i(x_0) \colon \lambda_i \geqslant 0 \; \forall i \text{ with } \lambda_i = 0 \text{ when } i \text{ inactive} \right\}$$

then $S_0 \subseteq \mathrm{NC}[A; x_0]$. We say that x_0 is *quasi-regular* if the opposite inclusion holds, that is if each outer normal is of the form $\sum \lambda_i \nabla g_i(x_0)$. It is known that x_0 is quasi-regular if either (a) each g_i is linear or (b) $\nabla g_i(x_0), \ldots, \nabla g_m(x_0)$ are linearly independent. See, for example, Hestenes [6, page 221 and page 241]. In the sequel, the above is generalized to the Banach space setting. Let then X, Y be Banach spaces over $\mathbb{R}$ and suppose that Y is partially ordered by a closed convex cone B. We are concerned to minimize a real-valued function f on X subject to $g(x) \in B$, where g is a given function from X into Y. We shall assume that g is Fréchet differentiable. Let $A = g^{-1}(B)$, and let $x_0 \in A$. Then one defines $\mathrm{TC}[A, x_0]$ as before, so that $\mathrm{NC}[A; x_0]$, the polar of $\mathrm{TC}[A; x_0]$, is a subset of the Banach dual space X^*. Let

$$S_1 = \{ y^* \circ g'(x_0) \colon y^* \in Y^*, y^*(g(x_0)) = 0, y^* \leqslant 0 \text{ on } B \}.$$

In the classical situation described at the beginning, if one lets $Y = \mathbb{R}^m$ and $B = \mathbb{R}^m_-$ (the natural negative cone) with $g(x) = (g_1(x), \ldots, g_m(x))$, then $S_0 = S_1$, and it is possible to define quasi-regularity as before. However, the objection might be raised that this definition should depend on the range $g(X)$, rather than on all of Y. Accordingly, we let

$$S_2 = \{ y^* \circ g'(x_0) \colon y^* \in Y^*, y^*(g(x_0)) = 0, y^* \leqslant 0 \text{ on } B \cap g(X) \}$$

Certainly $S_1 \subseteq S_2 \subseteq \mathrm{NC}[A, x_0]$, but examples will show that each inclusion can be proper.

DEFINITION. We say that x_0 is a *quasi-regular point* if $S_2 = \mathrm{NC}[A, x_0]$.

In the classical situation outlined previously, it is clear that $S_0 = S_1 = S_2$. Thus Theorems 1 and 2 do generalize the classical results stated in (a), (b) above. More generally, suppose Y has the property that, for every closed subspace M, every $m^* \in M^*$ satisfying $m^* \leqslant 0$ on $B \cap M$ can be extended to a functional $y^* \in Y^*$ satisfying $y^* \leqslant 0$ on B. Then again $S_1 = S_2$, and the distinction between the two possible definitions vanishes. This will be the case if Y is finite dimensional and B is a polyhedral cone. For some discussion of the "positive extension problem", we refer the reader to Asimow [1] and Hustad [7].

THEOREM 1. *Let Y be general (not necessarily finite dimensional). If $G = g'(x_0)$ maps onto Y, then x_0 is quasi-regular.*

PROOF. Let $y_0 = g(x_0)$, and let $f(x) = g(x + x_0) - y_0$. Then $f(0) = 0$, and $f'(0)$ is onto, so Lemma 1 of Flett [5] is applicable. This states that $f'(0)^{-1}(Q) \subseteq \mathrm{TC}[f^{-1}(Q); 0]$ whenever Q is a subset of X which is closed under multiplication by positive scalars. Examination of the proof shows that this is also true whenever Q is assumed merely to be closed under multiplication by scalars in $[0, 1]$. So we may put $Q = B - y_0$ and obtain $G^{-1}(B - y_0) \subseteq \mathrm{TC}[A; x_0]$. (Bender [2, Lemma 4.1] gives a simpler proof of a similar result but assumes that B has an interior point.) Now

$$B \subseteq B - y_0 \quad \text{and} \quad 0, \pm y_0 \in B - y_0,$$

and so

$$G^{-1}(B), G^{-1}(0), \pm G^{-1}(y_0) \subseteq G^{-1}(B - y_0) \subseteq \mathrm{TC}[A; x_0].$$

Hence, the polars

$$\left(G^{-1}(B)\right)^{\pi}, \left(G^{-1}(0)\right)^{\pi}, \left(\pm G^{-1}(y_0)\right)^{\pi} \supseteq \mathrm{NC}[A; x_0].$$

If x^* is in the intersection of the three polar sets on the left, then x^* vanishes on $G^{-1}(0)$ and so must be of the form $x^* = y^* \circ G$ for some $y^* \in Y^*$. Moreover, if $b \in B$, and if $x \in G^{-1}(b)$, then $x^*(x) \leqslant 0$, i.e. $y^*(b) \leqslant 0$. Similarly one can show that $y^*(y_0) = 0$.

THEOREM 2. *Let g be linear with $g(X)$ closed in Y. Then x_0 is quasi-regular.*

PROOF. Write Z for $g(X)$. Then $g'(x_0) = g$ maps onto the Banach space Z, and the admissible set A is unchanged when Y is replaced by Z. By Theorem 1, each outer normal x^* of A at x_0 is of the form $z^* \circ g$ with $z^* \in Z^*$, $z^*(y_0) = 0$ and $z^* \leqslant 0$ on $B \cap Z$. By the usual Hahn-Banach theorem there exists $y^* \in Y^*$ such that $y^* = z^*$ on Z. Thus $x^* = y^* \circ g$, i.e. $x^* \in S_2$.

The interest of quasi-regularity lies in the following generalization of Fritz John's theorem on the Lagrange multiplier rule.

THEOREM 3. *Let $f\colon X \to \mathbb{R}$ and $g\colon X \to Y$ be Fréchet differentiable functions, and let x_0 be a local minimum point for f subject to the constraint $g(x) \in B$. Suppose further that x_0 is quasi-regular. Then there exists $y^* \in Y^*$ with $y^*(g(x_0)) = 0$ and $y^*(b) \leqslant 0$ for all $b \in B \cap g(X)$ such that $f'(x_0) + y^* \circ g'(x_0) = 0$.*

PROOF. Let $A = g^{-1}(B)$ as before. We need only show that $-f'(x_0) \in \mathrm{NC}[A; x_0]$. Let $h \in \mathrm{TC}[A; x_0]$ be a unit vector, and take a sequence $\{a_q\}$ in A converging to x_0 such that $(a_q - x_0)\|a_q - x_0\|^{-1} \to h$. Then

$$\left(f(a_q) - f(x_0)\right)\|a_q - x_0\|^{-1} \to f'(x_0)h,$$

the value of $f'(x_0)$ at h. Since $f(x_0) \leqslant f(a_q)$ for all large q, it follows that $f'(x_0)h \geqslant 0$.

Of course the literature already contains a variety of Fritz John type theorems. They use assumptions and techniques different from ours to establish the existence of suitable Lagrange multipliers. For example Craven [3, Theorem 1] uses the hypothesis of "local solvability" [4, page 33] instead of our quasi-regularity. The sufficient conditions for quasi-regularity in Theorems 1 and 2 are also sufficient conditions for local solvability. However, the relationship between quasi-regularity and local solvability remains unclear. An advantage of our approach is the weakness of the definition of quasi-regularity, which allows Theorem 3 to hold for a wide variety of cones.

Our first example shows that S_1 can be a proper subset of S_2, and that S_2 is the most natural set to use in defining quasi-regularity.

EXAMPLE 1. Let $X = \mathbb{R}^2$, and let $Y = \mathbb{R}^3$ be ordered by the circular cone $B = \{(x, y, z)\colon z \geqslant \sqrt{x^2 + y^2}\}$. Define $g\colon X \to Y$ by $g(x, y) = (x, y, x)$, and let $x_0 = (0, 0)$. Clearly the feasible set A is $g^{-1}(B) = \{(x, 0)\colon x \geqslant 0\}$, and $B \cap g(X) = g(A) = \{(x, 0, x)\colon x \geqslant 0\}$. It follows easily that $\mathrm{TC}[A; x_0] = A$, and so $\mathrm{NC}[A; x_0]$ is the closed left half-plane $\{(x, y)\colon x \leqslant 0\}$. (Naturally we identify X^* with $\mathbb{R}^2$.) Since g is linear, we have $g'(x_0) = g$, and so $S_2 = \{y^* \circ g\colon y^*(1, 0, 1) \leqslant 0\}$ is also the closed left half-plane. Thus x_0 is quasi-regular. However

$$\begin{aligned} S_1 &= \{y^* \circ g\colon y^*(\cos\theta, \sin\theta, 1) \leqslant 0 \text{ for all } \theta\} \\ &= \left\{(x + z, y)\colon z \leqslant -\sqrt{x^2 + y^2}\right\} \\ &= \{(x, y)\colon x < 0\} \cup \{(0, 0)\}, \end{aligned}$$

which is not even a closed set.

We remark that circular cones have also served as counterexamples for other purposes [4, page 25].

Since the range $g(X)$ is generally not a linear subspace of Y, whereas $g'(x_0)$ is always a linear map, it might be considered natural to examine the following set:

$$S_3 = \{y^* \circ g'(x_0)\colon y^* \in Y^*,\ y^*(g(x_0)) = 0,\ y^* \leqslant 0 \text{ on } B \cap g'(x_0)(X)\}.$$

When g is linear, it is clear that $S_3 = S_2$. But even in the classical situation, S_3 need not be a subset of $\mathrm{NC}[A; x_0]$.

EXAMPLE 2. Let $X = \mathbb{R}$, and give $Y = \mathbb{R}^2$ the standard cone $B = \mathbb{R}^2_-$. Let $x_0 = 0$, and define $g\colon X \to Y$ by $g(x) = (-1 - x, x)$. Routine calculations show that $S_1 = S_2 = \mathrm{NC}[A; x_0] = \mathbb{R}_+$. Since $B \cap g'(x_0)(X) = \{0\}$, it follows that $S_3 = \mathbb{R}$. Thus S_3 is not a very natural set to work with in this context.

https://doi.org/10.1017/S1446788700033607 Published online by Cambridge University Press

Some examples of points which are not quasi-regular are given by Hestenes [6, pages 222–223].

Acknowledgement

The second author's work was supported by a Queen Elizabeth Fellowship.

References

[1] L. Asimow, 'Monotone extensions in ordered Banach spaces', *J. London Math. Soc.* **6** (1973), 563–569.
[2] P. J. Bender, 'Nonlinear programming in normed linear spaces', *J. Optim. Theory Appl.* **24** (1978), 263–285.
[3] B. D. Craven, 'Lagrangean conditions and quasiduality', *Bull. Austral. Math. Soc.* **16** (1977), 325–339.
[4] B. D. Craven, *Mathematical programming and control theory*, (Chapman and Hall, London, 1978).
[5] T. M. Flett, 'On differentiation in normed vector spaces', *J. London Math. Soc.* **42** (1967), 523–533.
[6] M. R. Hestenes, *Optimization theory, the finite dimensional case*, (Wiley, New York, 1975).
[7] O. Hustad, 'Linear inequalities and positive extension of linear functionals', *Math. Scand.* **8** (1960), 333–338.

Department of Mathematics
The Chinese University of Hong Kong
Shatin, N. T.
Hong Kong

Department of Mathematics
Institute of Advanced Studies
Australian National University
Canberra ACT 2601
Australia

https://doi.org/10.1017/S1446788700033607 Published online by Cambridge University Press