

AN INVERSE-FUNCTION THEOREM FOR A CLASS OF MULTIVALUED FUNCTIONS

STEPHEN M. ROBINSON¹

ABSTRACT. The inverse-function theorem is generalized to multivalued functions of the form $f(x)+K$, where f is a differentiable single-valued function and K is a nonempty closed convex cone. An application to Pareto optimization is given.

1. Introduction. The classical inverse-function theorem says that if the derivative of a differentiable function f between two Banach spaces is invertible at (hence near) a point x_0 , then the mapping itself is invertible near x_0 . In cases where the derivative at x_0 is onto, but not necessarily invertible, a similar result may be established via the implicit-function theorem [2, Theorem 1(4.XVII)]; in this case f will carry any neighborhood of x_0 onto a neighborhood of $f(x_0)$. In this paper we extend the latter result to the case of multivalued functions having the form $f(x)+K$, where f is a differentiable single-valued function and K is a nonempty closed convex cone. The results cited above then appear as the special case in which K consists of the origin alone. Multivalued functions of the form just mentioned are of practical interest in problems of optimization; for example, in §4 we show how to apply our main result to obtain a very simple derivation of the general necessary optimality conditions for Pareto optimization.

2. A lemma on contraction mappings. Multivalued contraction mappings have been studied by Nadler [5] and by Covitz and Nadler [1]. However, their results concerned only mappings which were either everywhere contractive or else uniformly locally contractive. We prove here a result about mappings which may only be contractive in a certain ball. We also obtain an error bound which will be useful in what follows.

DEFINITION. *In a metric space (X, ρ) the distance from a point x to a nonempty set $S \subset X$ is*

$$d(x, S) := \inf\{\rho(x, y) \mid y \in S\}.$$

Received by the editors November 6, 1972 and, in revised form, February 23, 1973.
AMS (MOS) subject classifications (1970). Primary 47H15; Secondary 49E30.

¹ Sponsored by the United States Army under Contract No. DA-31-124-ARO-D-462.

© American Mathematical Society 1973

The Hausdorff distance $\rho(A, B)$ between the nonempty sets A and B is

$$\rho(A, B) := \max\{\sup\{d(x, B) \mid x \in A\}, \sup\{d(y, A) \mid y \in B\}\}.$$

LEMMA. Let (X, ρ) be a complete metric space, and let T be a function from X into the class of subsets of X . Suppose there are nonnegative real numbers α and r , with $0 < \alpha < 1$, and a point $x_0 \in X$ such that:

(a) For some $\varepsilon > 0$ and all x, y in the closed ball $\bar{B}(x_0, r + \varepsilon)$ of radius $r + \varepsilon$ about x_0 , Tx and Ty are nonempty closed sets with $\rho(Tx, Ty) \leq \alpha\rho(x, y)$, and

(b) $d(x_0, Tx_0) \leq (1 - \alpha)r$.

Then there is a point $x_\infty \in \bar{B}(x_0, r + \varepsilon)$ with $x_\infty \in Tx_\infty$; that is, x_∞ is a fixed point of T . Further, $\rho(x_0, x_\infty) \leq (1 - \alpha)^{-1}d(x_0, Tx_0) + \varepsilon$.

PROOF. Denote $d(x_0, Tx_0)$ by δ . Choose a positive number γ with $\gamma\alpha(1 - \alpha)^{-1} \leq \varepsilon$, and find some $x_1 \in Tx_0$ with $\rho(x_0, x_1) \leq \delta + \frac{1}{2}\gamma\alpha \leq (1 - \alpha)r + \frac{1}{2}\gamma\alpha \leq r + \varepsilon$; thus $x_1 \in \bar{B}(x_0, r + \varepsilon)$. Suppose that for some $k \geq 1$ we have $x_{k-1} \in \bar{B}(x_0, r + \varepsilon)$ and $x_k \in Tx_{k-1}$ with

$$(1) \quad \rho(x_{k-1}, x_k) \leq \alpha^{k-1}\delta + (1 - 2^{-k})\gamma\alpha^k$$

and

$$(2) \quad \rho(x_0, x_k) \leq (1 - \alpha^k)r + \gamma \sum_{j=1}^k (1 - 2^{-j})\alpha^j.$$

Then $\rho(x_0, x_k) \leq r + \varepsilon$, so Tx_k is nonempty, and we can find $x_{k+1} \in Tx_k$ with $\rho(x_k, x_{k+1}) \leq d(x_k, Tx_k) + \gamma 2^{-(k+1)}\alpha^{k+1}$. Since

$$d(x_k, Tx_k) \leq \rho(Tx_{k-1}, Tx_k) \leq \alpha\rho(x_{k-1}, x_k),$$

we have $\rho(x_k, x_{k+1}) \leq \alpha^k\delta + (1 - 2^{-(k+1)})\gamma\alpha^{k+1}$. Also, $\rho(x_0, x_{k+1}) \leq \rho(x_0, x_k) + \rho(x_k, x_{k+1}) \leq (1 - \alpha^{k+1})r + \gamma \sum_{j=1}^{k+1} (1 - 2^{-j})\alpha^j$. Hence by induction the sequence $\{x_k\}$ exists, remains in $\bar{B}(x_0, r + \varepsilon)$ and satisfies (1) and (2) for each k . For $k \geq 0$ and any positive m , we have

$$(3) \quad \begin{aligned} \rho(x_k, x_{k+m}) &\leq \sum_{j=0}^{m-1} \rho(x_{k+j}, x_{k+j+1}) \leq \sum_{j=0}^{m-1} [\alpha^{k+j}\delta + (1 - 2^{-(k+j+1)})\gamma\alpha^{k+j+1}] \\ &\leq \alpha^k(1 - \alpha)^{-1}[\delta + \gamma\alpha], \end{aligned}$$

and so $\{x_k\}$ is a Cauchy sequence. Since (X, ρ) is complete, the sequence $\{x_k\}$ converges to some $x_\infty \in \bar{B}(x_0, r + \varepsilon)$. The bound for $\rho(x_0, x_\infty)$ follows from (3) by setting $k=0$ and recalling that $\gamma\alpha(1 - \alpha)^{-1} \leq \varepsilon$. To see that x_∞ is a fixed point, choose $k \geq 0$ and $\eta > 0$ and let $y \in Tx_\infty$ be a point with $\rho(x_{k+1}, y) \leq d(x_{k+1}, Tx_\infty) + \eta$. Then

$$d(x_\infty, Tx_\infty) \leq \rho(x_\infty, y) \leq \rho(x_\infty, x_{k+1}) + \rho(x_{k+1}, y).$$

But

$$\rho(x_{k+1}, y) \leq d(x_{k+1}, Tx_\infty) + \eta \leq \rho(Tx_k, Tx_\infty) + \eta \leq \alpha\rho(x_k, x_\infty) + \eta.$$

Hence $d(x_\infty, Tx_\infty) \leq \rho(x_\infty, x_{k+1}) + \alpha\rho(x_k, x_\infty) + \eta$, and this expression can be made as small as we wish by choosing k large and η small enough. Thus $d(x_\infty, Tx_\infty) = 0$, and since Tx_∞ is closed we have $x_\infty \in Tx_\infty$, which completes the proof.

If the sets Tx admit metric projections we can let $\varepsilon = \gamma = 0$, which simplifies the analysis. In this case, we obtain the point x_k by projecting x_{k-1} on Tx_{k-1} , and we have $\rho(x_k, x_\infty) \leq \alpha^k(1-\alpha)^{-1}d(x_0, Tx_0)$ for each k .

3. Generalization of the inverse-function theorem. In the following theorem, we employ certain results about *convex processes*: these are multivalued mappings between two linear spaces whose graphs are convex cones containing the origin. A convex process is thus a generalization of a linear transformation; if the graph is closed, we speak of a *closed convex process*. If T is a convex process from X into Y , its *domain*, $\text{dom } T$, is $\{x \mid Tx \neq \emptyset\}$; its *inverse* is another convex process defined on Y by $T^{-1}y := \{x \mid y \in Tx\}$. If X and Y are normed, we define the *norm* of T by

$$\|T\| := \sup\{\inf\{\|y\| \mid y \in Tx\} \mid \|x\| \leq 1, x \in \text{dom } T\}.$$

For more information about convex processes, see ([6], [7], [8]). In what follows we use S_X and S_Y to denote the closed unit balls in two Banach spaces X and Y . The closed ball of radius α about a point $a \in X$ is then denoted by $a + \alpha S_X$.

THEOREM 1. *Let X and Y be real Banach spaces, and let $K \subset Y$ be a nonempty closed convex cone. Let f be a Fréchet differentiable function from an open set $H \subset X$ into Y , and let its derivative f' be continuous at a point $x_0 \in H$. Define a multivalued function $F(\cdot)$ on H by $F(x) := f(x) + K$, and a convex process $F'(x_0)(\cdot)$ on all of X by*

$$F'(x_0)(x) := f'(x_0)x + K.$$

Suppose $F'(x_0)$ carries X onto Y , and let β be any number with $0 < \beta < 1$. Let θ be a positive number such that $x_0 + \theta S_X \subset H$ and such that for any $x \in x_0 + \theta S_X$, $\|F'(x_0)^{-1}\| \|f'(x) - f'(x_0)\| \leq \beta$. Let $\delta > 0$ be small enough so that $\|F'(x_0)^{-1}\| \delta \leq (1 - \beta)\theta$. Then for any α with $0 \leq \alpha < 1$, the image under F of the closed ball $x_0 + \alpha(2 - \alpha)\theta S_X$ contains $F(x_0) + \alpha\delta S_Y$.

PROOF. If $\alpha = 0$ the theorem is trivial; therefore suppose $\alpha > 0$. Choose any point $y \in F(x_0) + \alpha\delta S_Y$ and a point $y_0 \in F(x_0)$ with $\|y - y_0\| \leq \alpha\delta$.

Define a single-valued function $g(w)$ for $w \in H - x_0$ by $g(w) := f(x_0 + w) - f'(x_0)w$, and a multivalued function T_y from $H - x_0$ into X by $T_y w := F'(x_0)^{-1}[y - g(w)]$. By assumption the domain of $F'(x_0)^{-1}$ is all of Y , so $T_y w \neq \emptyset$ for each $w \in H - x_0$. Also, the graph of $F'(x_0)^{-1}$ is easily shown to be closed, so by [6, Corollary to Theorem 2], $\|F'(x_0)^{-1}\|$ is finite. Now let w_1 and w_2 be any two points in θS_X . By [6, Theorem 6] applied to $F'(x_0)^{-1}$ and the singleton sets $\{y - g(w_1)\}$ and $\{y - g(w_2)\}$ we have

$$(4) \quad \rho(T_y w_1, T_y w_2) \leq \|F'(x_0)^{-1}\| \|g(w_1) - g(w_2)\|.$$

Also, by [2, Chapter XVII, §1], we have

$$(5) \quad \begin{aligned} & \|F'(x_0)^{-1}\| \|g(w_1) - g(w_2)\| \\ & \leq \|w_1 - w_2\| \|F'(x_0)^{-1}\| \sup\{\|g'(\lambda w_1 + (1 - \lambda)w_2)\| \mid 0 < \lambda < 1\} \\ & \leq \beta \|w_1 - w_2\|. \end{aligned}$$

By definition,

$$\begin{aligned} d(0, T_y 0) &= \inf\{\|z\| \mid z \in F'(x_0)^{-1}[y - f(x_0)]\} \\ &= \inf\{\|z\| \mid y - f(x_0) \in f'(x_0)z + K\}. \end{aligned}$$

Since K is a convex cone, we have $[y_0 - f(x_0)] + K \subset K$, and thus

$$(6) \quad \begin{aligned} d(0, T_y 0) &\leq \inf\{\|z\| \mid y - f(x_0) \in f'(x_0)z + [y_0 - f(x_0)] + K\} \\ &= \inf\{\|z\| \mid y - y_0 \in f'(x_0)z + K\} \\ &= \inf\{\|z\| \mid z \in F'(x_0)^{-1}[y - y_0]\} \\ &\leq \|F'(x_0)^{-1}\| \|y - y_0\| \leq (1 - \beta)\alpha\theta, \end{aligned}$$

where we have used the definition of $\|F'(x_0)^{-1}\|$. We now apply the lemma with $r = \alpha\theta$, $\varepsilon = \alpha(1 - \alpha)\theta < (1 - \alpha)\theta$, contraction constant β , and starting point $w_0 = 0$. Using (4), (5), and (6) we conclude that there is a fixed point w_∞ of T_y with $\|w_\infty\| = \|w_\infty - w_0\| \leq (1 - \beta)^{-1}(1 - \beta)\alpha\theta + \alpha(1 - \alpha)\theta = \alpha(2 - \alpha)\theta$. Since $w_\infty \in T_y w_\infty = F'(x_0)^{-1}[y - g(w_\infty)]$, we have

$$y - [f(x_0 + w_\infty) - f'(x_0)w_\infty] \in f'(x_0)w_\infty + K;$$

that is, $y \in F(x_0 + w_\infty)$. But $x_0 + w_\infty \in x_0 + \alpha(2 - \alpha)\theta S_X$, which completes the proof.

If we define the *inverse* of F by $F^{-1}(y) := \{x \mid y \in F(x)\}$, then Theorem 1 says in particular that the domain of F^{-1} (that is, $\{y \mid F^{-1}(y) \neq \emptyset\}$) includes an open neighborhood of $F(x_0)$. A further analogy to the classical inverse-function theorem is provided by the fact that the convex process $F'(x_0)$ can reasonably be regarded as a Fréchet derivative of F at x_0 , since

$$\lim_{\|h\| \rightarrow 0} \|h\|^{-1} \rho(F(x_0 + h), F(x_0) + F'(x_0)h) = 0.$$

4. An application to Pareto optimization. In this section we illustrate an application of Theorem 1, using it to obtain a very short proof of the necessary optimality conditions for Pareto optimization. We recall that if U is a topological space, Λ is a subset of U , and θ is a function from U into a linear space W containing a nonempty convex cone P , then \bar{x} is said to be a *local Pareto optimizer* of θ on Λ with respect to P if $\bar{x} \in \Lambda$ and there is some open neighborhood $N(\bar{x})$ such that for no $x \in \Lambda \cap N(\bar{x})$ do we have $\theta(\bar{x}) \in \theta(x) + P \setminus \{0\}$. This concept reduces to that of ordinary local minimization if $W = \mathbf{R}$ and $P = \mathbf{R}_+$, the nonnegative real numbers. Finally, we say the cone P is *pointed* if $P \cap -P = \{0\}$ and *nontrivial* if $P \neq \{0\}$.

It is possible to use Theorem 1 to prove optimality conditions for problems in which the set Λ is defined by an infinite number of constraints, or by operator inequalities in partially-ordered spaces. However, in such cases one has to deal with a number of technical complications involving closure and complementarity, which detract from the simplicity of the argument. Since our purpose here is not to investigate optimality conditions in their utmost generality, but rather to illustrate an application of Theorem 1, we prefer to confine ourselves to a finite number of inequality and equality constraints, and to a finite-dimensional (but not necessarily polyhedral) cone P .

THEOREM 2. *Let X be a real Banach space, and let θ, g and h be Fréchet differentiable functions from an open set $H \subset X$ into $\mathbf{R}^l, \mathbf{R}^m$ and \mathbf{R}^a respectively. Let $P \subset \mathbf{R}^l$ be a nontrivial closed pointed convex cone. Suppose x_0 is a local Pareto optimizer of θ with respect to P on the set $\{x \in H \mid g(x) \leq 0, h(x) = 0\}$, and that $\theta', g',$ and h' are continuous at x_0 . Then there exist vectors $p \in P^*, u \in \mathbf{R}_+^m$ and $v \in \mathbf{R}^a$, not all zero, such that*

$$\theta'(x_0)*p = g'(x_0)*u + h'(x_0)*v$$

and $\langle u, g(x_0) \rangle = 0$, where

$$P^* := \{x \in \mathbf{R}^l \mid \langle x, p \rangle \leq 0 \text{ for each } p \in P\}.$$

Here and in what follows we denote by \mathbf{R}_+^m the nonnegative orthant of \mathbf{R}^m , and by \leq the usual partial ordering induced by this cone.

PROOF. Let $g_A(x)$ be a function from H into $\mathbf{R}^n, n \leq m$, composed of those components $g_i(x)$ for which $g_i(x_0) = 0$; that is, $g_A(x)$ is the vector of constraints active at x_0 . It follows immediately that x_0 is a local Pareto optimizer of θ with respect to P on the set $\{x \in H \mid g_A(x) \leq 0, h(x) = 0\}$. Now define a multivalued function $F: H \rightarrow \mathbf{R}^{l+n+a}$ by

$$F(x) := \begin{bmatrix} \theta(x) \\ g_A(x) \\ h(x) \end{bmatrix} + \begin{bmatrix} P \\ \mathbf{R}_+^n \\ \{0\} \end{bmatrix}.$$

By hypothesis, the point $[\theta(x_0)^T, 0, 0]^T \in \mathbf{R}^l \times \mathbf{R}^n \times \mathbf{R}^q$ lies in $F(x_0)$, where T denotes a transpose vector. However, there exists an open neighborhood $N(x_0) \subset H$ such that for arbitrarily small $p \in P \setminus \{0\}$, the point

$$[(\theta(x_0) - p)^T, 0, 0]^T$$

does not lie in the image under F of $N(x_0)$. For if it did lie in $F(x)$ for some $x \in N(x_0)$, then we should have $g_A(x) \leq 0$, $h(x) = 0$, and $\theta(x_0) - p \in \theta(x) + P$, so that $\theta(x_0) \in \theta(x) + p + P$. But since P is pointed, we have $p + P \subset P \setminus \{0\}$, so $\theta(x_0) \in \theta(x) + P \setminus \{0\}$, contradicting the local optimality of x_0 . Applying Theorem 1, we conclude that the convex process S defined by

$$Sx := \begin{bmatrix} \theta'(x_0)x \\ g'_A(x_0)x \\ h'(x_0)x \end{bmatrix} + \begin{bmatrix} P \\ \mathbf{R}_+^n \\ \{0\} \end{bmatrix}$$

does not carry X onto \mathbf{R}^{l+n+q} . By a standard hyperplane separation theorem [4, Theorem 3.2.3], we can separate the convex set $S(X)$ from a point not contained in it. Thus there exist a nonzero vector

$$[-p^T, u_A^T, v^T]^T \in \mathbf{R}^{l+n+q}$$

and a real number η such that for each $x \in X$, each $r \in P$ and each non-negative $w \in \mathbf{R}^n$,

$$\langle -p, \theta'(x_0)x + r \rangle + \langle u_A, g'_A(x_0)x + w \rangle + \langle v, h'(x_0)x \rangle \geq \eta.$$

By varying $r \in P$, $w \geq 0$ and $x \in X$ we obtain $p \in P^*$ and $u_A \geq 0$. Also, for $r=0$, $w=0$ and any $x \in X$ we have

$$(7) \quad \langle \theta'(x_0)^* (-p) + g'_A(x_0)^* u_A + h'(x_0)^* v, x \rangle \geq \eta,$$

which is possible only if the first element in the brackets is zero.

Now let $u \in \mathbf{R}_+^m$ be a vector whose components are equal to those of u_A for indices corresponding to constraints active at x_0 , and zero otherwise. Obviously $\langle u, g(x_0) \rangle = 0$, and from (7) we find

$$\theta'(x_0)^* p = g'(x_0)^* u + h'(x_0)^* v,$$

which completes the proof.

The usual constrained minimization problem of nonlinear programming can be obtained as a special case of the above by letting $P = \mathbf{R}_+$. For this problem, Theorem 2 establishes the fundamental necessary optimality conditions of Fritz John type (cf. [4, Theorem 11.3.1]), which generalize the classical Lagrange-multiplier conditions for equality-constrained problems.

In order to obtain optimality conditions of Kuhn-Tucker type (*i.e.*, with $p \neq 0$), one generally imposes a constraint qualification on the functions g_A and h . A suitable qualification in the present case is a requirement that the convex process taking x into

$$\begin{bmatrix} g'_A(x_0)x \\ h'(x_0)x \end{bmatrix} + \begin{bmatrix} \mathbf{R}_+^n \\ \{0\} \end{bmatrix}$$

carry X onto \mathbf{R}^{n+a} : indeed, the "modified Arrow-Hurwicz-Uzawa constraint qualification" of Mangasarian and Fromovitz [4, 11.3.5] is precisely this requirement.

We note that our treatment of the Pareto optimization problem (when specialized to the case $P = \mathbf{R}_+^l$ and with no equality constraints) differs in one important respect from the treatment given in [3]. There, some Pareto optima (those termed *improper* by Kuhn and Tucker) are excluded from consideration: in our notation, these correspond to Pareto optima x_0 such that

$$0 \in \begin{bmatrix} \theta'(x_0) \\ g'_A(x_0) \\ h'(x_0) \end{bmatrix} (X) + \begin{bmatrix} P \setminus \{0\} \\ \mathbf{R}_+^n \\ \{0\} \end{bmatrix}.$$

Such optima can be handled without difficulty by the method given here. However, the vectors of multipliers p associated with such improper optima will not in general lie in the interior of P^* , even if a constraint qualification is imposed, whereas the multipliers obtained in [3] (by excluding improper optima) do have this property.

ACKNOWLEDGEMENT. I am indebted to J.-P. Aubin for pointing out that the analysis of §4 can be carried out as well for Pareto optimization as for conventional minimization problems.

REFERENCES

1. H. Covitz and S. B. Nadler, Jr., *Multi-valued contraction mappings in generalized metric spaces*, Israel J. Math. **8** (1970), 5-11. MR **41** #7667.
2. L. V. Kantorovič and G. P. Akilov, *Functional analysis in normed spaces*, Fizmatgiz, Moscow, 1959; English transl., Internat. Series of Monographs in Pure and Appl. Math., vol. 46, Macmillan, New York, 1964. MR **22** #9837; **35** #4699.
3. H. W. Kuhn and A. W. Tucker, *Nonlinear programming*, Proc. Second Berkeley Sympos. on Math. Statist. and Probability, 1950, Univ. of California Press, Berkeley, Calif., 1951, pp. 481-492. MR **13**, 855.
4. O. L. Mangasarian, *Nonlinear programming*, McGraw-Hill, New York, 1969. MR **40** #5263.
5. S. B. Nadler, Jr., *Multi-valued contraction mappings*, Pacific J. Math. **30** (1969), 475-488. MR **40** #8035.

6. S. M. Robinson, *Normed convex processes*, Trans. Amer. Math. Soc. **174** (1972), 127-140.

7. R. T. Rockafellar, *Monotone processes of convex and concave type*, Mem. Amer. Math. Soc. No. 77 (1967). MR **37** #825.

8. ———, *Convex analysis*, Princeton Math. Series, no. 28, Princeton Univ. Press, Princeton, N.J., 1970. MR **43** #445.

MATHEMATICS RESEARCH CENTER, UNIVERSITY OF WISCONSIN, MADISON, WISCONSIN
53706