A SECOND-ORDER EXTENSION OF LJUSTERNIK'S THEOREM WITHOUT TWICE FRECHET DIFFERENTIABILITY CONDITION

Furukawa, Nagata
Department of Mathematics, Faculty of Science, Kyushu University

https://doi.org/10.5109/13422

出版情報：Bulletin of informatics and cybernetics. 25 (1/2), pp.53-59, 1992-03. Research Association of Statistical Sciences

Bulletin of Informatics and Cybernetics, Vol. 25, No. 1~2, 1992

A SECOND-ORDER EXTENSION OF LJUSTERNIK'S THEOREM WITHOUT TWICE FRÉCHET DIFFERENTIABILITY CONDITION

By

Nagata FURUKAWA*

Abstract

For a mapping f from a Banach space to another space, a second-order extension of Ljusternik's theorem is given under certain assumptions weaker than those of previous literatures. Twice Fréchet differentiability of f is not assumed. In place of it, the second-order Neustadt derivative and directional derivative are used.

1. Introduction

A second-order version of Lujusternik's theorem has been given, for instance on a finite dimensional space, by Ben-Israel, Ben-Tal and Zlobec (Ref. 1, Theorem 12.9) as follows:

Let $H: R^n \to R^p$ be twice continuously differentiable at a point x_0 satisfying $H(x_0) = 0$. Assume that the range space of the Jacobian matrix $\nabla H(x_0)$ is R^p. Let $V_H(x_0, x_1)$ be defined by

$$V_H(x_0, x_1) = \left\{ z \in R^n \;\middle|\; \begin{array}{l} \exists\, t_0 > 0,\ \exists r: (0, t_0] \to R^n \\ \textit{such that} \\ x_0 + t x_1 + \frac{1}{2} t^2 z + r(t) \in N[H]\ \forall t \in (0, t_0) \\ \|r(t)\|/t^2 \to 0 \textit{ as } t \downarrow 0 \end{array} \right\},$$

where $N[H] = \{x \in R^n;\ H(x) = 0\}$. Then for every x_1 satisfying $\nabla H(x_0) x_1 = 0$, it holds that

$$V_H(x_0, x_1) = \{z \in R^n;\ \nabla H(x_0)\, z + H''(x_0)(x_1, x_1) = 0\}, \tag{1}$$

where $H''(x_0)$ is the Hessian matrix of H at x_0.

A similar result on Banach spaces has been given by Ben-Tal and Zowe (Ref. 2, Proposition 7.2).

The purpose of this paper is to give an expression similar to (1), under weakened assumptions concerning the twice differentiability of the mapping H. In this paper H is a mapping from a Banach space to another space, and the twice Fréchet differentiability of the mapping is not imposed, while the first-order Fréchet differentiability is assumed.

---

* Department of Mathematics, Faculty of Science, Kyushu University, Fukuoka, Japan

In place of the twice Fréchet differentiability, we use the Neustadt derivative and one-sided (curved) directional derivative, both of which are of the second-order. In Section 3 of this paper, an expression of second-order tangent direction for the kernel of the mapping is given in terms of second-order directional derivative of the mapping.

2. Basic Concepts

First we shall define two kinds of second-order variational sets.

DEFINITION 2.1. *Let X be a real Banach space, and let Q an arbitrary subset of X. For $x_0 \in clQ$ and $x_1 \in X$, two kinds of variational sets of Q at x_0 (with respect to x_1) are defined by the following:*

$$V(Q;\, x_0,\, x_1) = \left\{ h \in X \;\middle|\; \begin{array}{l} \exists\, t_0 > 0,\ \exists r:(0,\, t_0] \to X \\ \textit{such that} \\ x_0 + tx_1 + t^2h + r(t) \in Q \textit{ for } \forall t \in (0,\, t_0] \\ \|r(t)\|/t^2 \to 0 \textit{ as } t \downarrow 0 \end{array} \right\},$$

$$T(Q;\, x_0,\, x_1) = \left\{ h \in X \;\middle|\; \begin{array}{l} \exists\, \{y_n\} \subset Q,\ \exists\, \{\lambda_n\} \downarrow 0 \\ \textit{such that} \\ (y_n - x_0 - \lambda_n x_1)/\lambda_n^2 \to h \textit{ as } n \to \infty \end{array} \right\}.$$

$V(Q;\, x_0,\, x_1)$ is an extension of the tangent cone to the second-order case, and $T(Q;\, x_0,\, x_1)$ a second-order extension of the cone of adherent displacements. These sets are the special cases of higher-order variational sets given by Furukawa and Yoshinaga (Ref. 4). We should note that the definition of $V(Q;\, x_0,\, x_1)$ is slightly different from that of $V_H(x_0,\, x_1)$ in the coefficient of t^2, when we put $Q = N[H]$.

Next we shall introduce the concepts of second-order Neustadt derivative and one-sided (curved) directional derivative.

DEFINITION 2.2. *Let X and Y be real Banach spaces, and f be a mapping from X to Y. Let x_0 be a point of X. Suppose that f has the first-order Fréchet derivative at x_0, say $Df(x_0)$. If, for a given point $x_1 \in X$, there exists a point $f^{(2)}(x_0,\, x_1;\, x)$ of Y such that*

$$f^{(2)}(x_0,\, x_1;\, x) = \lim_{\substack{y \to x \\ \lambda \downarrow 0}} \frac{1}{\lambda^2}[f(x_0 + \lambda x_1 + \lambda^2 y) - f(x_0) - \lambda Df(x_0)x_1]$$

for every $x \in X$, then the mapping $x \mapsto f^{(2)}(x_0,\, x_1;\, x)$ is called the second-order Neustadt derivative of f at x_0 with respect to x_1. In this case f is said to be twice Neustadt differentiable at x_0 with respect to x_1.

DEFINITION 2.3. *Let f be as in Definition 2.2. If, for a given point $x_1 \in X$, there exists a point $f''(x_0,\, x_1;\, x)$ of Y such that*

$$f''(x_0,\, x_1;\, x) = \lim_{\lambda \downarrow 0} \frac{1}{\lambda^2}[f(x_0 + \lambda x_1 + \lambda^2 x) - f(x_0) - \lambda Df(x_0)x_1]$$

for every $x \in X$, *then we call the mapping* $x \mapsto f''(x_0, x_1, x)$ *the second-order directional derivative of* f *at* x_0 *with respect to* x_1.

The following definition is essentially same as the notion of local Lipschitz.

DEFINITION 2.4. *(Ref. 3). Let* $f: X \to Y$ *where* X *and* Y *are Banach spaces, and let* $x_0 \in X$. *When there exist a neighborhood* U *of* x_0 *and a constant* K *such that*

$$\|f(x') - f(x'')\|_Y \leq K \|x' - x''\|_X \text{ for } \forall x', \forall x'' \in U,$$

f *is said to be Lipschitz near* x_0.

The following proposition is obvious from the definitions.

PROPOSITION 2.1. *Assume that* $f: X \to Y$ *is Lipschitz near* x_0. *If there exists the second-order Neustadt derivative* $f^{(2)}(x_0, x_1;\cdot)$, *then there exists the second-order directional derivative* $f''(x_0, x_1;\cdot)$ *and one has* $f^{(2)}(x_0, x_1;\cdot) = f''(x_0, x_1;\cdot)$. *The converse relation is also true.*

3. Main Results

Let X and Y be a real Banach spaces. Throughout this section we shall use the following notation for the kernel of a mapping $g: X \to Y$, that is,

$$N[g] = \{x \in X \mid g(x) = 0\}.$$

Now let f be a mapping from X into Y. We shall explain some inclusion relations among two kinds of second-order variational sets for the kernel $N[f]$ and the kernel set of the second-order Neustadt derivative of f in what follows.

LEMMA 3.1. *Let* $x_0 \in N[f]$. *Assume that* f *is Fréchet differentiable at* x_0, *and twice Neustadt differentiable at* x_0 *with respect to* x_1 *for every* $x_1 \in N[Df(x_0)]$. *Then the following results hold for all* $x_1 \in N[Df(x_0)]$:

(i) $v(N[f]; x_0, x_1) \subset N(f^{(2)}(x_0, x_1;\cdot)]$,

(ii) $T(N[f]; x_0, x_1) \subset N[f^{(2)}(x_0, x_1;\cdot)]$,

where $N[f^{(2)}(x_0, x_1;\cdot)]$ *denotes the kernel of the mapping* $x \mapsto f^{(2)}(x_0, x_1; x)$.

PROOF. Let $x_1 \in N[Df(x_0)]$. We first prove (i).

Let h be any point of $V(N[f]; x_0, x_1)$. By the definition, then, there exist a positive number t_0 and a mapping $r: (0, t_0] \to X$ such that

$$x_0 + tx_1 + t^2 h + r(t) \in N[f] \quad \text{for } \forall t \in (0, t_0], \tag{2}$$

$$\|r(t)\|/t^2 \to 0 \text{ as } t \downarrow 0. \tag{3}$$

We put

$$\begin{aligned} \phi(t) &= f(x_0 + tx_1 + t^2 h + r(t)) - f(x_0) - t\, Df(x_0)x_1 \\ &= f(x_0 + tx_1 + t^2(h + r(t)/t^2)) - f(x_0) - t\, Df(x_0)x_1. \end{aligned}$$

Since $h + r(t)/t^2 \to h$ as $t \downarrow 0$ by (3), from the definition of the second-order Neustadt derivative we have

$$\phi(t)/t^2 \to f^{(2)}(x_0, x_1; h) \text{ as } t \downarrow 0. \tag{4}$$

Obviously it follows from (2) and the way of taking x_0 and x_1 that

$$\phi(t) \equiv 0 \text{ on } (0, t_0].$$

Hence from (4) it follows that $f^{(2)}(x_0, x_1; h) = 0$, that is, $h \in N[f^{(2)}(x_0, x_1; \cdot)]$.

Next we shall prove (ii). Let $h \in T(N[f]; x_0, x_1)$. Then there exist $\{y_n\} \subset N[f]$ and $\{\lambda_n\} \downarrow 0$ such that

$$(y_n - x_0 - \lambda_n x_1)/\lambda_n^2 \to h \text{ as } n \to \infty. \tag{5}$$

Hence one has

$$[f(y_n) - f(x_0) - \lambda_n Df(x_0)x_1]/\lambda_n^2 = 0 \text{ for } \forall n,$$

which yields with the help of (5) that

$$\begin{aligned} 0 &= \lim_{n\to\infty}[f(y_n) - f(x_0) - \lambda_n Df(x_0)x_1]/\lambda_n^2 \\ &= \lim_{n\to\infty}[f(x_0 + \lambda_n x_1 + \lambda_n^2(y_n - x_0 - \lambda_n x_1)/\lambda_n^2) - f(x_0) - \lambda_n Df(x_0)x_1]/\lambda_n^2 \\ &= f^{(2)}(x_0, x_1; h). \end{aligned}$$

Thus $h \in N[f^{(2)}(x_0, x_1; \cdot)]$. This completes the proof.

The converse inclusion relations to (i) and (ii) in the above lemma are very important. To derive the converse relations we need to set up the following assumption.

ASSUMPTION H. (i) *f is Fréchet differentiable* (*in the sense of the first order*) *on a neighborhood U of* x_0, (ii) *the Fréchet derivative* $Df(\cdot)$ *is continuous as a mapping from U to* $L(X, Y)$ (*the space of continuous linear mappings from X to Y endowed with the operator norm topology*), *and* (iii) $Df(x_0)X = Y$.

The following proposition is an easy consequence of Corollary in p. 32 of Ref. 3.

PROPOSITION 3.1. *Suppose that f satisfies* (i) *and* (ii) *in Assumption H. Then f is Lipschitz near* x_0.

We are now ready to prove our main theorem.

THEOREM 3.1. *Let* $x_0 \in N[f]$. *Let the assumptions of Lemma* 3.1 *be satisfied, and let Assumption H be satisfied. Then it holds that*

$$V(N[f]; x_0, x_1) = T(N[f]; x_0, x_1) = N[f^{(2)}(x_0, x_1; \cdot)] = N[f''(x_0, x_1; \cdot)] \tag{6}$$

for all $x_1 \in N[Df(x_0)]$.

PROOF. Let $x_1 \in N[Df(x_0)]$. Since, under Assumption H, f is Lipschitz near x_0 by Proposition 3.1, Proposition 2.1 implies that the last two sets in (6) are equal. Hence, by virtue of Lemma 3.1, it suffices to show the following two relations:

$$N[f^{(2)}(x_0, x_1;\cdot)] \subset V(N[f]; x_0, x_1), \tag{7}$$

$$N[f^{(2)}(x_0, x_1;\cdot)] \subset T(N[f]; x_0, x_1). \tag{8}$$

Let $h \in N[f^{(2)}(x_0, x_1;\cdot)]$. From generalized Ljusternik's theorem given by Ioffe and Tihomirov (Ref. 5, p. 34), there exist a neighborhood U of x_0, a constant $K > 0$ and a mapping $\tau: U \to X$ such that

$$f(\xi + \tau(\xi)) = 0 \qquad \forall \xi \in U, \tag{9}$$

$$\|\tau(\xi)\| \leq K\,\|f(\xi) - f(x_0)\| = K\|f(\xi)\| \qquad \forall \xi \in U. \tag{10}$$

In the right-hand side of (10), notice that $x_0 \in N[f]$. We set

$$v(t) = x_0 + tx_1 + t^2h,$$

then we can take a positive small number t_0 such that

$$v(t) \in U \qquad \forall t \in [0, t_0]. \tag{11}$$

We have

$$\limsup_{t\downarrow 0} \frac{1}{t^2}\|\tau(v(t))\| \leq \limsup_{t\downarrow 0} \frac{K}{t^2}\|f(v(t))\| \qquad \text{(by (10) and (11))}$$

$$= K \limsup_{t\downarrow 0} \frac{1}{t^2}\|f(x_0 + tx_1 + t^2h)\| \tag{12}$$

$$= K \limsup_{t\downarrow 0} \frac{1}{t^2}\|f(x_0) + t\,Df(x_0)x_1 + t^2\cdot\frac{1}{t^2}[f(x_0 + tx_1 + t^2h) - f(x_0) - t\,Df(x_0)\,x_1]\|.$$

Since f is Lipschitz near x_0 as mentioned above, it holds that

$$[f(x_0 + tx_1 + t^2h) - f(x_0) - t\,Df(x_0)\,x_1]/t^2 \to f^{(2)}(x_0, x_1; h) \text{ as } t \downarrow 0.$$

Thus the last term of (12) is equal to

$$\text{K} \limsup_{t\downarrow 0} \frac{1}{t^2}\,\|f(x_0) + t\,Df(x_0)x_1 + t^2 f^{(2)}(x_0, x_1; h)\,\| = 0, \tag{13}$$

in which the last equality follows from that $x_0 \in N[f]$, $x_1 \in N[Df(x_0)]$ and $h \in N[f^{(2)}(x_0, x_1;\cdot)]$. Combining (12) with (13) yields

$$\lim_{t\downarrow 0} \|r(t)\|/t^2 = 0, \tag{14}$$

where

$$r(t) = \tau(v(t)).$$

Substituting $\xi = v(t)$ in (9) yields, by virtue of (9) and (11),

$$f(v(t) + r(t)) = 0 \;\forall t \in [0, t_0].$$

Namely

$$f(x_0 + tx_1 + t^2h + r(t)) = 0 \ \forall t \in [0, t_0],$$

which implies, together with (14), that

$$h \in V(N[f]; x_0, x_1).$$

This completes the proof of (7).

Let again $h \in N[f^{(2)}(x_0, x_1;\cdot)]$. Let $\{\lambda_k\} \downarrow 0$ be arbitrary but fixed. Since $f^{(2)}(x_0, x_1;\cdot) = f''(x_0, x_1;\cdot)$ under Assumption H, we have

$$\begin{aligned}
&\lim_{k\to\infty} \frac{1}{\lambda_k^2} f(x_0 + \lambda_k x_1 + \lambda_k^2 h) \\
&= \lim_{k\to\infty} \frac{1}{\lambda_k^2}[f(x_0 + \lambda_k x_1 + \lambda_k^2 h) - f(x_0) - \lambda_k Df(x_0)x_1] \\
&= f''(x_0, x_1; h) = 0,
\end{aligned}$$

which implies

$$\lim_{k\to\infty} \frac{1}{\lambda_k^2} \|f(x_0 + \lambda_k x_1 + \lambda_k^2 h)\| = 0. \tag{15}$$

Set

$$\xi_k = x_0 + \lambda_k x_1 + \lambda_k^2 h,$$

then from (10) there exists a positive integer k_0 such that

$$\begin{aligned}
&\xi_k \in U \qquad \forall k \geq k_0, \\
&\|\tau(\xi_k)\| \leq K \, \|f(\xi_k)\| \qquad \forall k \geq k_0.
\end{aligned} \tag{16}$$

Hence, by virtue of (15) and (16),

$$\limsup_{k\to\infty} \frac{1}{\lambda_k^2} \|\tau(\xi_k)\| \leq \limsup_{k\to\infty} \frac{K}{\lambda_k^2} \|f(\xi_k)\| = 0,$$

which implies

$$\lim_{k\to\infty} \frac{1}{\lambda_k^2} \|\tau(\xi_k)\| = 0. \tag{17}$$

On the other hand we get from (9)

$$f(\xi_k + \tau(\xi_k)) = 0 \qquad \forall k \geq k_0.$$

Therefore, letting $z_k = \xi_k + \tau(\xi_k)$, we have $\{z_k;\, k \geq k_0\} \subset N[f]$. Moreover,

$$\begin{aligned}
(z_k - x_0 - \lambda_k x_1)/\lambda_k^2 &= h + \tau(\xi_k)/\lambda_k^2 \\
&\to h \text{ as } k \to \infty,
\end{aligned}$$

by virtue of (17). These relations imply that

$$h \in T(N[f]; x_0, x_1),$$

which completes the proof.

COROLLARY 3.1. *Let $x_0 \in N[f]$. Assume that f is twice Fréchet differentiable at x_0, and let $D^2 f(x_0)$ be the second-order Fréchet derivative of f at x_0. Let Assumption H be satisfied. Then*

$$\begin{aligned} V(N[f]; x_0, x_1) &= T(N[f]; x_0, x_1) \\ &= \{x \in X;\ Df(x_0)x + \tfrac{1}{2}D^2 f(x_0)\,(x_1, x_1) = 0\}. \end{aligned} \tag{18}$$

PROOF. Since f is twice Fréchet differentiable at x_0 by the assumptions, f is twice Neustadt differentiable at x_0 with respect to every x_1 and

$$f^{(2)}(x_0, x_1; x) = Df(x_0)x + \tfrac{1}{2}D^2 f(x_0)\,(x_1, x_1) \qquad \forall x \in X.$$

Hence (18) is obvious from Theorem 3.1.

REMARK. One will find a difference between the two expression formulae in (1) and (18): the coefficients of the second-order derivatives are not equal. This is due to the formal difference between the definitions of second-order tangent directions in Ref. 1 and in our present paper.

References

[1] BEN-ISRAEL, A., BEN-TAL, A. and ZLOBEC, S.: *Optimality in Nonlinear Programming: A Feasible Directions Approach*, John Wiley and Sons, 1981.

[2] BEN-TAL, A. and ZOWE, J.: *A Unified Theory of First and Second Order Conditions for Extremum Problem in Topological Vector Spaces*, Mathematical Programming Study **19**, (1982), 39–76.

[3] CLARKE, F.H.: *Optimization and Nonsmooth Analysis*, John Wiley and Sons, 1983.

[4] FURUKAWA, N. and YOSHINAGA, Y.: *Higher-order Variational Sets, Variational Derivatives and Higher-order Necessary Conditions in Abstract Mathematical Programming*, Bull. Informatics and Cybernetics, **23**, (1988), 9–40.

[5] IOFFE, A.D. and TIHOMIROV, V.M.: *Theory of Extremal Problems*, North-Holland Pub. Comp., 1979.

Received September 28, 1991