Techniques of Variational Analysis

by

Jonathan Borwein and Qiji Zhu

Addenda and Errata

Last updated November 29, 2013

Contents

1 Lemma 5.1.11: May 5, 2006

We thank John Vanderwerff for pointing out that a compactness assumption is missing in Lemma 5.1.11. Related corrections are:

1. page 169 line -2: add 'When Y is compact,' at the beginning of the sentence.
2. page 170 line 1: X should be Y.
3. page 170 line 6: change to: 'and if in addition Y is compact, then'
4. page 174 line 15: add before the word 'and': 'with Y compact'

2 Lemma 5.5.4: July 21, 2006

In the proof of Lemma 5.5.4 on page 232 line 24: Theorem 5.5.2 should be replaced by the following Theorem A (Theorem 2.7 in [174]). Also, on page 233 the second line from bottom: the second x^* should be p^*. We thank Professor W. Schirotzek for brought our attention to these inaccuracies.

Theorem A. (Subdifferential of Marginal Functions) *Let X and Y be Fréchet smooth Banach spaces, let $\phi(\cdot,\cdot) : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a lsc function and define $f(x) := \inf_{y\in Y} \phi(x,y)$. Suppose that $x^* \in \partial_F \underline{f}(x)$ where $\underline{f}$ is the lsc closure of f. Then, for any $\varepsilon > 0$, there exist $(x_\varepsilon, y_\varepsilon)$ and there exists an element of the joint Fréchet subdifferential of ϕ, $(x^*_\varepsilon, y^*_\varepsilon) \in \partial_F \phi(x_\varepsilon, y_\varepsilon)$ such that $\|x - x_\varepsilon\| < \varepsilon$, $|\underline{f}(x) - \underline{f}(x_\varepsilon)| < \varepsilon$,*

$$\phi(x_\varepsilon, y_\varepsilon) < \underline{f}(x_\varepsilon) + \varepsilon < f(x_\varepsilon) + \varepsilon, \tag{1}$$

and

$$\|x^*_\varepsilon - \xi\| < \varepsilon, \ \ \|y^*_\varepsilon\| < \varepsilon. \tag{2}$$

Proof. Let g be a C^1 function such that $g'(x) = x^*$ and $\underline{f} - g$ attains a minimum 0 at x over $B_r(x)$ for some $r \in (0,1)$. Let $\varepsilon \in (0,r)$ be an arbitrary positive number. Choose $\eta < \min(\varepsilon/5, \sqrt{\varepsilon/10})$ such that $x' \in B_\eta(x)$ implies that $\underline{f}(x') \geq \underline{f}(x) - \varepsilon/5$, $\|g'(x') - g'(x)\| < \varepsilon/2$ and $\|x' - x''\| < \eta$ implies that $|g(x') - g(x'')| < \varepsilon/5$. Taking $z \in B_{\eta/2}(x)$ close enough to x so that $f(z) - g(z) < \underline{f}(x) - g(x) + \eta^2/16$, one can choose $y \in Y$ satisfying

$$\begin{aligned} \phi(z,y) - g(z) &< f(z) - g(z) + \eta^2/16 \\ &< \underline{f}(x) - g(x) + \eta^2/8 \\ &\leq \inf_{(u,v)\in B_r(x)\times Y} (\phi(u,v) - g(u)) + \eta^2/8. \end{aligned} \tag{3}$$

Applying the Smooth Variational Principle of Theorem 3.1.10 with $p = 2$, $\lambda = 1$ and $\varepsilon = \eta^2/8$ to the function $(u,v) \to \phi(u,v) - g(u)$ we obtain $(x_\varepsilon, y_\varepsilon) \in B_{\eta/2}(z,y) \subset B_\eta(x,y)$ and a C^1 function h such that

$$\max(\|h\|_\infty, \|h'\|_\infty) < \eta,$$

and the function

$$(u,v) \to \phi(u,v) - g(u) + h(u,v)$$

attains its minimum at $(u,v) = (x_\varepsilon, y_\varepsilon)$. Then

$$(g'(x_\varepsilon) - h'_x(x_\varepsilon, y_\varepsilon), -h'_y(x_\varepsilon, y_\varepsilon)) \in \partial_F \phi(x_\varepsilon, y_\varepsilon)$$

and (2) follows with $x^*_\varepsilon = g'(x_\varepsilon) - h'_x(x_\varepsilon, y_\varepsilon)$ and $y^*_\varepsilon = -h'_y(x_\varepsilon, y_\varepsilon)$. It remains to verify inequality (1) and $|\underline{f}(x) - \underline{f}(x_\varepsilon)| < \varepsilon$:

$$\begin{aligned} \phi(x_\varepsilon, y_\varepsilon) &\leq \varphi(z,y) + [g(x_\varepsilon) - g(z)] + h(z,y) - h(x_\varepsilon, y_\varepsilon) \\ &\leq \underline{f}(x) + 2\eta^2 + |g(x_\varepsilon) - g(z)| + 2\|h\|_\infty \\ &< \underline{f}(x_\varepsilon) + \varepsilon/5 + 2\eta^2 + |g(x_\varepsilon) - g(z)| + 2\|h\|_\infty < f(x_\varepsilon) + \varepsilon. \end{aligned}$$

Since $\underline{f}(x_\varepsilon) \leq f(x_\varepsilon) \leq \phi(x_\varepsilon, y_\varepsilon)$ we have $|\underline{f}(x) - \underline{f}(x_\varepsilon)| < \varepsilon$. ●

3 Theorem 3.3.8: August 18, 2006

Theorem 3.3.8 should be a corollary of a stronger version of Theorem 3.3.7 with the condition, for $n = 1, \ldots, N$,

$$\liminf_{x \to \bar{x}} d(\partial_F f_n(x), 0) > 0 \tag{4}$$

replaced by the weaker condition

$$\liminf_{(x,f_n(x))\to(\bar{x},f(\bar{x}))} d(\partial_F f_n(x),0) > 0. \tag{5}$$

We thank Professor W. Schirotzek for pointing this out. Related changes are listed below:

- Replace condition (4) by the weaker condition (5) in Theorems 3.3.4 and 3.3.7. and in the second line from below on page 63.
- On page 60 line 2. After the first sentence add: Moreover, $g_i(y_i) \leq 0$ and $y_i \to \bar{x}$ force $f(y_i) \in (f(\bar{x})-\varepsilon/2, f(\bar{x})+\varepsilon/2)$ for i sufficiently large.
- On page 60 line 4: add after 'such that': $|f(x_i) - f(\bar{x})| < \varepsilon$ and

4 Theorems 2.1.1 and 2.1.4: November 1, 2006

Conclusion (iii) in Theorem 2.1.1 (page 7 line 4) should be

(iii) $f(x) + \varepsilon d(x,y) > f(y)$, for all $x \in X\backslash\{y\}$.

Correspondingly, the last two sentences of the proof (page 8 lines 7 and 8) should be changed to "Thus, $f(x) + \varepsilon d(x,z_i) > f(z_i) \geq f(y) + \varepsilon d(y,z_i)$. Combining this with the triangle inequality we arrive at (iii)."

The qualification "for all $x \in X\backslash\{y\}$" should also be added to the conclusion of Theorem 2.1.4 on page 9 line 10. We thank Professor H. Bauschke for alerting us to these corrections.

5 Theorem 3.7.2: November 22, 2006

We thank Professor W. Schirotzek for brought our attention to the need of showing $s_1(x) > 0$ in the proof of Theorem 3.7.2. A corrected proof is given below.

Theorem 3.7.2 *Let X be a Fréchet smooth Banach space and let $M_n, n = 1,2,\ldots,N$ be metric spaces. Consider closed-valued multifunctions $S_n\colon M_n \to$*

X, $n = 1, 2, \ldots, N$. Suppose that $\bar{x}$ is an extremal point of the extremal system $(S_1, S_2, \ldots, S_N)$ at $(\bar{m}_1, \bar{m}_2, \ldots, \bar{m}_N)$. Then for any $\varepsilon > 0$, there exist $m_n \in B_\varepsilon(\bar{m}_n), x_n \in B_\varepsilon(\bar{x}), n = 1, 2, \ldots, N$ and $x_n^* \in N_F(S_n(m_n); x_n) + \varepsilon B_{X^*}, n = 1, 2, \ldots, N$ such that $\max\{\|x_n^*\| \mid n = 1, \ldots, N\} \geq 1$ and

$$x_1^* + x_2^* + \cdots + x_N^* = 0.$$

Proof. Let U be a neighborhood of $\bar{x}$ as in the definition of an extremal point. Without loss of generality we may assume that $U = B_\varepsilon(\bar{x})$. Choose $\varepsilon' \in (0, \varepsilon/2)$ satisfying

$$(4N^2 + 1)\varepsilon' + 4N(\varepsilon')^2 < \varepsilon^2/64$$

and let $m_1, m_2, \ldots, m_N$ be as in the definition of the extremal point for $\varepsilon = \varepsilon'$. Let s_1 be as in Lemma 3.2.2 and define

$$f_1(y_1, y_2, \ldots, y_N) := s_1(y_1, y_2, \ldots, y_N),$$

$$f_2(y_1, y_2, \ldots, y_N) := \sum_{n=1}^{N} \iota_{S_n(m_n)}(y_n),$$

and

$$r(y_1, \ldots, y_N) := \sum_{n=1}^{N} \left\| y_n - \bar{x} \right\|^2.$$

Choose $y_n' \in S_n(m_n), n = 1, 2, \ldots, N$ such that $\|y_n' - \bar{x}\| < d(S_n(m_n); \bar{x}) + \varepsilon' < 2\varepsilon'$. Then

$$\bigwedge[f_1 + f_2, r](X^N) \leq (f_1 + f_2 + r)(y_1', y_2', \ldots, y_N') < 4N^2\varepsilon' + 4N(\varepsilon')^2.$$

Applying the nonlocal approximate sum rule of Theorem 3.2.3 we have that there exist $u = (u_1, \ldots, u_N)$, $v = (v_1, \ldots, v_N)$, and $u^* = (u_1^*, \ldots, u_N^*) \in \partial_F(f_1 + f_2)(u)$ such that

$$\|u - v\| < \varepsilon', \tag{6}$$

$$(f_1 + f_2)(u) + r(v) < \bigwedge[f_1 + f_2, r](X^N) + \varepsilon' < (4N^2 + 1)\varepsilon' + 4N(\varepsilon')^2, \tag{7}$$

and

$$\|u^* - r'(v)\| < \varepsilon' < \frac{\varepsilon}{4}. \tag{8}$$

Note that (7) implies

$$r(v) = \sum_{n=1}^{N} \left\|v_n - \bar{x}\right\|^2 < (4N^2+1)\varepsilon' + 4N(\varepsilon')^2. \tag{9}$$

Consequently,

$$\|r'(v)\| < 2N\sqrt{(4N^2+1)\varepsilon' + N(\varepsilon')^2} < \varepsilon/4, \tag{10}$$

and, therefore,

$$\|u^*\| < \varepsilon/2. \tag{11}$$

The inequality (7) also implies that $u_n \in S_n(m_n)$ so that $f_1(u) = s_1(u) > 0$. Since $f_1 = s_1$ is Lipschtiz, applying the strong local approximate sum rule of Theorem 3.3.19 we can find $x = (x_1, \ldots, x_N)$ and $z = (z_1, \ldots, z_N)$ in $B_{\varepsilon/2}(u)$, $-x^* = -(x_1^*, \ldots, x_N^*) \in \partial_F f_1(z) = \partial_F s_1(z)$ and $z^* = (z_1^*, \ldots, z_N^*) \in \partial_F f_2(x)$ such that $s_1(z) > 0$, and

$$\|z^* - x^* - u^*\| < \frac{\varepsilon}{2}. \tag{12}$$

Since

$$\partial_F f_2(x) = N_F(S_1(m_1); x_1) \times \cdots \times N_F(S_N(m_N); x_N),$$

combining (11) and (12) we have

$$x_n^* \in N_F(S_n(m_n); x_n) + \varepsilon B_{X^*}.$$

Inequalities (6), (9) and $x \in B_{\varepsilon/2}(u)$ imply that $x_n \in B_\varepsilon(\bar{x})$. Finally, since $s_1(z) > 0$, it follows from Lemma 3.2.2 that $x_1^* + x_2^* + \cdots + x_N^* = 0$ and $\max\{\|x_n^*\| \mid n = 1, \ldots, N\} \geq 1$, which completes the proof. ●

6 Exercise 4.3.11. September 15, 2008

The set I should be $I(x) := \{i\colon g_i(x) = g(x)\}$ and the point must be assumed to be one at which all functions are continuous. In the generality given we needed to include normal cones to various domains. Indeed the result as stands is correct if $0\partial g_i(x)$ is interpreted as $\partial 0\, g_i(x) = N_{\operatorname{dom} g_i}(x)$.

7 Exercise 3.4.11. September 30, 2009

The exercise and the hint should be:

Exercise 3.4.11 Deduce the Lipschitz criterion from the cone monotonicity criterion. Hint: For each unit vector $u \in X$, consider the R_+u monotonicity of $f(\cdot) \pm L\langle u^*, \cdot\rangle$ where $u^* \in X^*$ is a unit vector satisfying $\langle u^*, u\rangle = 1$.

8 Section 4.7.3. November 18, 2009

Professor Henry Wolkowicz points out to us that in the fourth line below the section title, the definition of double stochastic pattern should be:

We say that *a nonnegative matrix* A has a *double stochastic pattern* if there is a doubly stochastic matrix with exactly the same zero entries as A.

9 Example 5.1.19, November 23, 2013

'upper semicontinuous' should be 'sequentially upper semicontinuous'. Related, in the first sentence of page 174 'a metric space' should be 'a compact metric space'.

Professor Kim Border point out to us that the lower level set of $f(x, y) = \max(x, y)$ is not upper semicontinuous in the usual sense that f^{-1} preserves open sets.

10 Example 4.5.13, November 29, 2013

D should be the closed convex cone not just the convex hull.

11 Section 1.2. May 12, 2015

Page 3 in the end of first paragraph "in its domain" should be eliminated. We thank Mrs. Truong Xuan Duc Ha from the Institute of Mmathematics, Hanoi, Vietnam for pointing this out.