Extremal Principle in Variational Analysis

January 24, 2009

1 (2.2.3) Extremal Characterizaions of Asplund Spaces

Definition 1 *(2.17 Asplund Spaces)*
A Banach Space X is Asplund (or has the Asplund property) if every convex continuous function $\varphi : U \to \mathbb{R}$ defined on an open convex subset U of X is Fréchet differentiable on a dense subset of U.

Theorem 2 *(2.20 extremal Characterizations of Asplund Spaces)*
Let X be a Banach space. TFAE

(a) *X is Asplund.*

(b) *The approximate extremal principle holds in X.*

(c) *The ε-extremal principle holds in X.*

Corollary 3 *(2.21 Boundary characterizaions of Asplund Spaces)*
Let X be a Banach space and $\Omega \subset X$ be a closed proper subset. TFAE

(a) *X is Asplund.*

(b) *The set $\left\{x \in \operatorname{bd}\Omega : \widehat{N}(x;\Omega) \neq 0\right\}$ is dense in $\operatorname{bd}\Omega$.*

(c) *There exists an $x \in \operatorname{bd}\Omega$ such that $\widehat{N}(x;\Omega) \neq 0$.*

(d) *For all $\varepsilon > 0$ and all $M \geq \varepsilon$, the set $\left\{x \in \operatorname{bd}\Omega : \widehat{N}_\varepsilon(x;\Omega) \backslash M\mathbf{B}^* \neq \phi\right\}$ is dense in $\operatorname{bd}\Omega$.*

(e) *For all $\varepsilon > 0$ and all $M \geq \varepsilon$ there exists an $x \in \operatorname{bd}\Omega$ such that $\widehat{N}_\varepsilon(x;\Omega) \backslash M\mathbf{B}^* \neq \phi$.*

Theorem 4 *(2.22 exact extremal principle in Asplund Spaces)*

(i) *Let X be an Asplund space and $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ be an extremal system in X, where all the Ω_is are closed around $\overline{x}$ and all but one of them are sequentially normally compact at $\overline{x}$. Then the exact extremal principle holds for $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$.*

(ii) *Conversely, let X be a Banach space such that the exact extremal principle holds for any extreremal system $\{\Omega_1, \Omega_2; \overline{x}\}$ where both sets are closed around $\overline{x}$ and one of them is sequentially normally compact at $\overline{x}$. Then X is Asplund.*

Corollary 5 *(2.24 Nontriaviality of the basic normals in Asplund spaces)*
Let X be an Asplund space and $\Omega \subset X$ be a closed proper and sequentially normally compact subset. Then $N(x; \Omega) \neq 0$ for all $x \in \operatorname{bd}\Omega$.

Corollary 6 *(2.25 subdifferentiability of Lipschitzian funcitons on Asplund spaces)*
Let X be a Banach space, $\overline{x} \in X$. Then for every function $\varphi : X \to \mathbb{R}$ locally Lipschitzian around $\overline{x}$, $\partial\varphi(\overline{x}) \neq 0$ if and only if X is Asplund.

2 (2.3) Relations with Variational Principles

- By classical variational priciples we mean a group of results stating that for any lsc function $\varphi : X \to \mathbb{R}$ bounded from below and a point x_0 close to the infimum one can find an arbitrarily small perturbation θ such that the perturbing function $\varphi + \theta$ acheives its minimum at a point $\overline{x}$ close to x_0.

2.1 (2.3.1) Ekland Variational Principle

Theorem 7 *(2.26 Ekland's variational principle)*
Let (X, d) be a metric space. Then

(i) *Suppose X is complete and $\varphi : X \to \overline{\mathbb{R}}$ is a proper l.s.c function which is bounded from below. Let $\varepsilon, \lambda > 0$ and $x_0 \in X$ be given such that*

$$\varphi(x_0) < \inf_X \varphi + \varepsilon. \tag{1}$$

. Then there exists an $\overline{x} \in X$ satisfying:

(a) $\varphi(\overline{x}) \leq \varphi(x_0)$,

(b) $d(\overline{x}, x_0) \leq \lambda$,

(c) $\varphi(x) + (\varepsilon/\lambda)\, d(\overline{x}, x) > \varphi(\overline{x})$ *for all* $x \neq \overline{x}$.

(ii) *Conversely, X is complete if for every Lipschitz continuous function $\varphi : X \to \mathbb{R}$ bounded from below and every $\varepsilon > 0$ there exists an $\overline{x} \in X$ satisfying*

(a') $\varphi(\overline{x}) \leq \inf_X \varphi + \varepsilon$ *and property (c) above.*

Proof. To show (i) it suffices to take $\lambda = \varepsilon = 1$, for otherwise, we may replace φ by $\varepsilon^{-1}\varphi$ and d by $\lambda^{-1}d$. Assume we are given $x_0 \in X$ such that (1) is satisfied. Let $\delta = \inf_X \varphi + 1 - \varphi(x_0)$. Inductively define the sequences $\{W_k\}_{k=1}^{\infty}, \{x_k\}_{k=1}^{\infty}$ as follows

$$W_k = \{x \in X : \varphi(x) + d(x, x_{k-1}) \leq \varphi(x_{k-1})\},$$

and choose $x_k \in W_k$ such that

$$\varphi(x_k) < \inf_{W_k} \varphi(x) + \frac{\delta}{2k}$$

Then W_k is nonempty (since $x_{k-1} \in W_k$) and it follows from the lower semicontinuity of φ that W_k is also closed (show this). Futhermore, the sequence $\{W_k\}_{k=1}^{\infty}$ is decreasing since if $x \in W_{k+1}$ then

$$\begin{aligned} \varphi(x) + d(x, x_{k-1}) &\leq \varphi(x) + d(x, x_k) + d(x_{k-1}, x_k) \\ &\leq \varphi(x_k) + d(x_k, x_{k-1}) \leq \varphi(x_{k-1}). \end{aligned}$$

Hence, $x \in W_k$. Moreover, for $x \in W_{k+1} \subset W_k$,

$$d(x, x_{k+1}) \leq \varphi(x_k) - \varphi(x) < \varphi(x) + \frac{1}{2k} - \varphi(x) = \frac{1}{2k}.$$

Therefore, $\operatorname{diam} W_{k+1} \leq \frac{\delta}{k} \to 0$ as $k \to \infty$. By the completness of X, $\bigcap W_k = \{\overline{x}\}$. For any $x \neq \overline{x}$, $x \notin W_k$ for sufficiently large k. Therefore, $\varphi(x) + d(x, x_k) > \varphi(x_k)$ for sufficiently large k. Taking the lim inf on both sides as $k \to \infty$ we get (c). To show (a) we note that, since $\overline{x}, x_0 \in W_1$,

$$\varphi(\overline{x}) + d(\overline{x}, x_0) \leq \varphi(x_0).$$

Therefore,

$$\varphi(\overline{x}) \leq \varphi(x_0).$$

To show (b), note first that $\overline{x}, x_1 \in W_2$. Then

$$\begin{aligned} d(\overline{x}, x_0) &\leq d(\overline{x}, x_1) + d(x_1, x_0) \\ &\leq \operatorname{diam} W_2 + \varphi(x_1) + d(x_1, x_0) - \varphi(x_1) \\ &\leq \delta + \varphi(x_0) - \varphi(x_1) \\ &= \inf_X \varphi + 1 - \varphi(x_0) + \varphi(x_0) - \varphi(x_1) \\ &\leq 1. \end{aligned}$$

To show (ii), let $\{x_k\}_{k=1}^{\infty}$ be a Cauchy sequence in X. Define the function $\varphi : X \to \mathbb{R}$ by

$$\varphi(x) := \lim_{k \to \infty} d(x, x_k).$$

- The limit exists since

  $$|d(x, x_k) - d(x, x_m)| \leq d(x_k, x_m) \to 0$$

  as $m, n \to \infty$.

- φ is Lipschitz continuous:

$$|d(x,x_k) - d(y,x_k)| \leq d(x,y).$$

Taking the limit as $k \to \infty$ we get

$$|\varphi(x) - \varphi(y)| \leq d(x,y).$$

- $\varphi(x_k) \to 0$ as $k \to \infty$: Let $\varepsilon > 0$ and choose K sufficiently large that $d(x_k, x_m) \leq \varepsilon$ for all $k, m \geq K$. Then, fixing $k \geq K$,

$$\varphi(x_k) = \lim_{m\to\infty} d(x_k, x_m) \leq \varepsilon.$$

Fix $0 < \varepsilon < 1$. By assumption, there is an $\overline{x} \in X$ such that $\varphi(\overline{x}) \leq \inf_X \varphi + \frac{\varepsilon}{2}$ and property (c) is satisfied with $\lambda = 1$.

- $\varphi(\overline{x}) = 0$:

$$\varphi(\overline{x}) < \varphi(x_k) + \varepsilon d(x_k, \overline{x}).$$

Since $\varphi(\overline{x}) := \lim_{k\to\infty} d(\overline{x}, x_k)$, $d(\overline{x}, x_k) \leq \varphi(\overline{x}) + \gamma$ for arbitrary $\gamma > 0$ and sufficiently large k. Therefore,

$$\varphi(\overline{x}) < \varphi(x_k) + \varepsilon(\varphi(\overline{x}) + \gamma).$$

Taking the limit as $k \to \infty$,

$$\varphi(\overline{x}) \leq \varepsilon(\varphi(\overline{x}) + \gamma).$$

Therefore,

$$\varphi(\overline{x}) \leq \frac{\varepsilon\gamma}{1-\varepsilon}.$$

Since γ is arbitrary, $\varphi(\overline{x}) = 0$, ie, $\lim_{k\to\infty} d(\overline{x}, x_k) = 0$.

■

Corollary 8 *(2.27 ε-stationary condition).*
Let X be a Banach space and $\varphi : X \to \mathbb{R}$ be proper, lsc and bounded from below. Given $\varepsilon, \lambda > 0$ and $x_0 \in X$ such that (1) is satisfied, assume that φ is Fréchet differentiable on a neighbourhood U of x_0 containing $B(x_0, \lambda)$. Then there exists an $\overline{x} \in X$ with $\|\overline{x} - x_0\| \leq \lambda$ such that $\varphi(\overline{x}) \leq \varphi(x_0)$ and $\|\varphi'(\overline{x})\| \leq \varepsilon/\lambda$.

Proof. By Theorem 2.26, there exists an $\overline{x} \in X$ such that $\varphi(\overline{x})$ is a local minimum of $\varphi(x) + \varepsilon/\lambda \|x - \overline{x}\|$. Therefore, $0 \in \partial(\varphi(x) + \varepsilon/\lambda \|x - \overline{x}\|)|_{x=\overline{x}} = \varphi'(\overline{x}) + \varepsilon/\lambda \mathbf{B}^*$. ■

2.2 (2.3.2) Subdifferential Variational Principle

Theorem 9 *(2.28: lower subdifferential variational principle)*
Let X be a Banach space. TFAE

(a) *The approximate extemal principle holds in X.*

(b) *For every proper lsc function $\varphi : X \to \overline{\mathbb{R}}$ bounded from below, every $\varepsilon > 0, \lambda > 0$ and $x_0 \in X$ with $\varphi(x_0) < \inf_X \varphi + \varepsilon$, there exist $\overline{x} \in X$ and $x^* \in \widehat{\partial}\varphi(\overline{x})$ such that $\|\overline{x} - x_0\| < \lambda$, $\varphi(\overline{x}) < \inf_X \varphi + \varepsilon$ and $\|x^*\| < \varepsilon/\lambda$.*

(c) *X is Asplund.*

Proof. It was shown before that (c) $\Longrightarrow$ (a).

(b) $\Longrightarrow$ (c): Let $\varphi : X \to \mathbb{R}$ be any convex continuous function. We will show that $\partial\varphi(x) \neq \phi$ for x in a dense subset of X. Since φ is convex, $\widehat{\partial}\varphi(x) \neq \phi$ for all $x \in X$. Therefore, it suffices to show that $\widehat{\partial}^{+}\varphi(x) \neq \phi$ for x in a dense subset of X. Fix $x_0 \in X$, $\varepsilon > 0$. Let $\psi := -\varphi$. Since ψ is continuous, there exists a $\nu > 0$ such that $\psi(x) > \psi(x_0) - \varepsilon$ for all $x \in x_0 + \nu\mathbf{B}$. Define the function $\theta : X \to \overline{\mathbb{R}}$ by

$$\theta(x) := \psi(x) + \delta(x, x_0 + \nu\mathbf{B}).$$

Then $\theta(x_0) < \inf_X \theta + \varepsilon$. Applying (b) we find an $\overline{x} \in X$ and $x^* \in \widehat{\partial}\theta(\overline{x})$ such that $\|\overline{x} - x_0\| < \nu$ and $\|x^*\| < \varepsilon/\nu$. Since $\overline{x} \in x_0 + \nu\mathbf{B}$, $x^* \in \widehat{\partial}\theta(\overline{x}) = \widehat{\partial}\psi(\overline{x}) = \widehat{\partial} - \varphi(\overline{x}) = \widehat{\partial}^{+}\varphi(\overline{x})$. Thus, $\partial\varphi(x) \neq \phi$ for x in a dense subset of X.

(a) $\Longrightarrow$ (b): Let $\varphi : X \to \overline{\mathbb{R}}$ be proper, lsc bounded from below. Given $\varepsilon > 0, \lambda > 0$ choose $\widetilde{\varepsilon}, \widetilde{\lambda}$ such that $\widetilde{\varepsilon} < \varepsilon$, $\widetilde{\lambda} < \lambda$, $\widetilde{\varepsilon}/\widetilde{\lambda} < \varepsilon/\lambda$ and let $x_0 \in X$ be such that $\varphi(x_0) < \inf_X \varphi + \widetilde{\varepsilon}$. By the Ekland variational principle (Theorem 2.26(i)), there exists an $\widetilde{x} \in X$, $\|\widetilde{x} - x_0\| < \widetilde{\lambda}$ such that $\varphi(\widetilde{x}) < \inf_X \varphi + \widetilde{\varepsilon}$ and

$$\varphi(\widetilde{x}) < \varphi(x) + \widetilde{\varepsilon}/\widetilde{\lambda}\,\|x - \widetilde{x}\| \;\; \forall x \neq x_0.$$

Define the two closed sets

$$\begin{aligned} \Omega_1 &: = \operatorname{epi}\varphi, \\ \Omega_2 &: = \operatorname{hypo}\left(\varphi(\widetilde{x}) - \widetilde{\varepsilon}/\widetilde{\lambda}\,\|x - \widetilde{x}\|\right). \end{aligned}$$

It is easy to check that the system $\{\Omega_1, \Omega_2; (\widetilde{x}, \varphi(\widetilde{x}))\}$ is an extremal system in $X \times \mathbb{R}$. By (a), for every $\widehat{\varepsilon} > 0$, there exists $(x_i^*, \xi_i) \in \widehat{N}((x_i, \alpha_i); \Omega_i)$, $i = 1, 2$ such that

$$\begin{cases} \|(x_i, \alpha_i) - (\widetilde{x}, \varphi(\widetilde{x}))\| < \widehat{\varepsilon}, \\ \frac{1}{2} - \widehat{\varepsilon} < \|(x_i^*, \xi_i)\| < \frac{1}{2} + \widehat{\varepsilon}, \\ \|(x_1^*, \xi_1) + (x_2^*, \xi_2)\| < \widehat{\varepsilon}. \end{cases}$$

Since, for sufficiently small $\widehat{\varepsilon}$, $(x_2^*, \xi_2) \neq 0$, $(x_2, \alpha_2) \in \operatorname{bd}\Omega_2$. Therefore, $(x_2, \alpha_2) = \left(x_2, \varphi(\widetilde{x}) - \widetilde{\varepsilon}/\widetilde{\lambda}\,\|x_2 - \widetilde{x}\|\right)$, which implies that $\xi_2 > 0$. (To see this, observe that $(x_2^*, \xi_2) \in$

$\widehat{N}((x_2,\varphi(x_2));\Omega_2)$ gives $\xi_2 \geq 0$. But if $\xi_2 = 0$ then $x_2^* = 0$ from the definitions of the Fréchet normals.) Furhermore, we get $-x_2^*/\xi_2 \in \widehat{\partial}^+ \left(\varphi(\widetilde{x}) - \widetilde{\varepsilon}/\widetilde{\lambda}\,\|\cdot - \widetilde{x}\|\right)(x_2) = \widetilde{\varepsilon}/\widetilde{\lambda}\mathbf{B}^*$, which gives

$$\|x_2^*\|/\xi_2 \leq \widetilde{\varepsilon}/\widetilde{\lambda}.$$

It follows from $\|(x_i^*,\xi_i)\| > \left(\frac{1}{2} - \widehat{\varepsilon}\right)$ that

$$\xi_2 \geq \frac{1/2 - \widehat{\varepsilon}}{\left(1 + \widetilde{\varepsilon}/\widetilde{\lambda}\right)}.$$

Since the expression on the right tends to $\frac{1}{2\left(1+\widetilde{\varepsilon}/\widetilde{\lambda}\right)} > 0$ as $\widehat{\varepsilon} \to 0$, it follows that $\xi_2 > \widehat{\varepsilon}$ for sufficiently small $\widehat{\varepsilon}$. This, toghether with $|\xi_1 + \xi_2| < \widehat{\varepsilon}$ give $\xi_1 < 0$. Since $(x_1^*,\xi_1) \in \widehat{N}((x_1,\alpha_1);\operatorname{epi}\varphi)$,

$$\overline{\lim_{(x,\alpha)\overset{\operatorname{epi}\varphi}{\to}(x_1,\alpha_1)}} \frac{\langle x_1^*, x - x_1\rangle + \xi_1(\alpha - \alpha_1)}{\|x - x_1\| + |\alpha - \alpha_1|} \leq 0.$$

In particular, if we take $x = x_1$,

$$\underline{\lim}_{\alpha \geq \varphi(x_1)} \frac{(\alpha - \alpha_1)}{|\alpha - \alpha_1|} \geq 0$$

which is possible only if $\alpha_1 = \varphi(x_1)$. Hence,

$$(x_1^*,\xi_1) \in \widehat{N}((x_1,\varphi(x_1));\operatorname{epi}\varphi)$$

and thus

$$x_1^*/|\xi_1| \in \widehat{\partial}\varphi(x_1)$$

and

$$\frac{\|x_1^*\|}{|\xi_1|} \leq \frac{\|x_2^*\| + \widehat{\varepsilon}}{\xi_2 - \widehat{\varepsilon}} \leq \frac{\widetilde{\varepsilon}/\widetilde{\lambda}\xi_2 + \widehat{\varepsilon}}{\xi_2 - \widehat{\varepsilon}} = \frac{\widetilde{\varepsilon}/\widetilde{\lambda} + \widehat{\varepsilon}/\xi_2}{1 - \widehat{\varepsilon}/\xi_2}.$$

Now the inequality $\xi_2 \geq \frac{1/2 - \widehat{\varepsilon}}{\left(1+\widetilde{\varepsilon}/\widetilde{\lambda}\right)}$ implies that $\widehat{\varepsilon}/\xi_2 \to 0$ as $\widehat{\varepsilon} \to 0$. This together with the expression above and the choice $\widetilde{\varepsilon}/\widetilde{\lambda} < \varepsilon/\lambda$ yeild , for sufficiently small $\widehat{\varepsilon}$,

$$\frac{\|x_1^*\|}{|\xi_1|} \leq \varepsilon/\lambda.$$

Also

$$\|x_1 - x_0\| \leq \|x_1 - \widetilde{x}\| + \|\widetilde{x} - x_0\| \leq \widehat{\varepsilon} + \widetilde{\lambda} \leq \lambda$$

for sufficiently small $\widehat{\varepsilon}$. Finally,

$$\varphi(x_1) = \alpha_1 \leq \varphi(\widetilde{x}) + \widehat{\varepsilon} < \inf_X \varphi + \widehat{\varepsilon} + \widetilde{\varepsilon} \leq \inf_X \varphi + \varepsilon$$

for sufficiently small $\widehat{\varepsilon}$. ■

Corollary 10 *(2.29 Fréchet subdifferentiability of lsc funcionctions)*
Let $\mathcal{A}$ be the class of all proper lsc funciotns $\varphi : X \to \overline{\mathbb{R}}$ on a Banach space X. The following peroperties are equivalent:

(a) *X is Asplund.*

(b) *For every $\varphi \in \mathcal{A}$ the set of points $\left\{(x, \varphi(x)) \in X \times \mathbb{R} : \widehat{\partial}\varphi(x) \neq \phi\right\}$ is dense in the graph of φ.*

(c) *For every $\varphi \in \mathcal{A}$ there exists an $x \in \operatorname{dom}\varphi$ with $\widehat{\partial}\varphi(x) \neq \phi$.*

(d) *For every $\varphi \in \mathcal{A}$ and every $\varepsilon > 0$ there exists an $x \in \operatorname{dom}\varphi$ with $\widehat{\partial}_{g\varepsilon}\varphi(x) \neq \phi$.*

(e) *For every $\varphi \in \mathcal{A}$ and every $\varepsilon > 0$ there exists an $x \in \operatorname{dom}\varphi$ with $\widehat{\partial}_{a\varepsilon}\varphi(x) \neq \phi$.*

Theorem 11 *(2.30: upper subdifferential variational principle)*
Let X be a Banach space and let $\varphi : X \to \mathbb{R}$ be a lsc function bounded from below. Then for every $\varepsilon > 0, \lambda > 0,$ and $x_0 \in X$ with $\varphi(x_0) < \inf_X \varphi + \varepsilon$ ther is $\overline{x} \in X$ with $\|\overline{x} - x_0\| < \lambda$ and $\varphi(\overline{x}) < \inf_X \varphi + \varepsilon$ such that

$$\|x^*\| < \varepsilon/\lambda$$

whenever $x^* \in \widehat{\partial}^+\varphi(\overline{x})$.