Extremal Principle in Variational Analysis

January 2, 2009

1 (2.1) Set Extremality and Nonconvex Separation

1.1 (2.1.1) Extremal Sequences of Sets

Definition 1 *(2.1 local extremality of systems of sets)*

Suppose $\Omega_1, \Omega_2, \ldots, \Omega_n$ *are nonempty subsets of a space* X, $n \geq 2$ *and* $\overline{x} \in \bigcap_{i=1}^{n} \Omega_i$. $\overline{x}$ *is called a local extremal point of the system of sets* $\Omega_1, \Omega_2, \ldots, \Omega_n$ *if given* $\varepsilon > 0$ *there exist elements* $a_1, a_2, \ldots, a_n \in \varepsilon\mathbf{B}$ *and a neighbourhood* U *of* $\overline{x}$ *such that*

$$\bigcap_{i=1}^{n} (\Omega_i - a_i) \cap U = \phi.$$

In this case the system $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ *is called an extremal system in* X.

- For $n = 2$, the local extremality of $\{\Omega_1, \Omega_2; \overline{x}\}$ can equivalently be described as follows: There is a neighbourhood U of $\overline{x}$ such that $\forall \varepsilon > 0$ there is an element $a \in \varepsilon\mathbf{B}$ such that
$$(\Omega_1 - a) \cap \Omega_2 \cap U = \phi.$$
- $\Omega_1 \cap \Omega_2 = \{\overline{x}\}$ does not imply that $\{\Omega_1, \Omega_2; \overline{x}\}$ is an extremal system; e.g., $\Omega_1 :=$ x-axis, $\Omega_2 := y$-axis.
- If $\overline{x}$ is a boundary point of Ω then $\{\Omega; \overline{x}\}$ is an extremal system.
- If the constrained optimization problem
$$\min \varphi(x) \ \ni \ x \in \Omega \subset X$$
has a solution $\overline{x}$ then $\{\operatorname{epi}\varphi, \Omega \times \{\varphi(\overline{x})\}; (\overline{x}, \varphi(\overline{x}))\}$ is an extremal system.

Proposition 2 *(2.2 interiors of sets in extremal systems)*

If $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ *is an extremal system then*

$$\bigcap_{i=1}^{n-1} \operatorname{int}(\Omega_i) \cap \Omega_n \cap U = \phi. \tag{1}$$

Proof. Suppose, to the contrary, that $u \in \bigcap_{i=1}^{n-1} \text{int}(\Omega_i) \cap \Omega_n \cap U$. Then there is an $\eta > 0$ such that $u + \eta \mathbf{B} \subset \bigcap_{i=1}^{n-1} \text{int}\,(\Omega_i) \cap U$. It follows that, for $i = 1, 2, \ldots, n-1$,

$$u + \frac{1}{2}\eta \mathbf{B} \subset \Omega_i - a \;\forall a \in \frac{1}{4}\eta \mathbf{B}.$$

Since $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ is an extremal system, there exist $a_1, a_2, \ldots, a_n \in \frac{1}{4}\eta \mathbf{B}$ such that $\bigcap_{i=1}^{n} (\Omega_i - a_i) \cap U = \phi$. Hence, $d\,(u, (\Omega_n - a_n)) \geq \frac{1}{2}\eta$ and $d\,(u, \Omega_n) \geq \frac{1}{4}\eta$ (because $d\,(u, (\Omega_n - a_n)) \leq d\,(u, \Omega_n) + d_H\,(\Omega_n, (\Omega_n - a_n)))$. This means that $u \notin \Omega_n$, which is a contradiction. ■

Definition 3 *(Separated sets)*

The sets $\Omega_1, \Omega_2, \ldots, \Omega_n$ are said to be separated if there exist functionals $x_1^*, x_2^*, \ldots, x_n^* \in X^*$, not all zeroes, and real numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$\begin{aligned} \langle x_i^*, x \rangle &\leq \alpha_i \;\forall x \in \Omega_i,\; i = 1, 2, \ldots, n \\ \sum_{i=1}^{n} x_i^* &= 0, \\ \sum_{i=1}^{n} \alpha_i &\leq 0. \end{aligned}$$

- If $\Omega_1, \Omega_2, \ldots, \Omega_n$ are separated and have common points then $\sum_{i=1}^{n} \alpha_i = 0$.

Proposition 4 *(2.3 extremality and separation)*

Let $\Omega_1, \Omega_2, \ldots, \Omega_n$ be nonempty subsets of a space X, $n \geq 2$ and $\overline{x} \in \bigcap_{i=1}^{n} \Omega_i$. Then

(i) *If $\Omega_1, \Omega_2, \ldots, \Omega_n$ are separated then $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ is an extremal system.*

(ii) *The converse holds if Ω_i are convex and* $\text{int}\,\Omega_i \neq \phi$, $i = 1, 2, \ldots, n-1$.

Proof. (i) Assume that $\Omega_1, \Omega_2, \ldots, \Omega_n$ are separated and $\overline{x} \in \bigcap_{i=1}^{n} \Omega_i$. Then there exist $x_1^*, x_2^*, \ldots, x_n^* \in X^*$, not all zeroes, and real numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$\begin{aligned} \langle x_i^*, x \rangle &\leq \alpha_i \;\forall x \in \Omega_i,\; i = 1, 2, \ldots, n \\ \sum_{i=1}^{n} x_i^* &= 0, \\ \sum_{i=1}^{n} \alpha_i &\leq 0. \end{aligned}$$

Assume, without loss of generality that $x_1^* \neq 0$. For all $\varepsilon > 0$ we can choose $a \in \varepsilon\mathbf{B}$ such that $\langle x_1^*, a \rangle > 0$. We are going to show that $(\Omega_1 - a) \cap \Omega_2, \ldots \cap \Omega_n = \phi$. Suppose, to the contrary, that $u \in (\Omega_1 - a) \cap \Omega_2, \ldots \cap \Omega_n$. Then

$$\langle x_1^*, u \rangle = \langle x_1^*, u + a \rangle - \langle x_1^*, a \rangle < \alpha_1.$$

On the other hand

$$\langle x_1^*, u \rangle = -\sum_{i=2}^{n} \langle x_i^*, u \rangle \geq -\sum_{i=2}^{n} \alpha_i \geq \alpha_1$$

which is a contradiction.

(ii) Suppose $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ is an extremal system, where Ω_i are convex and $\operatorname{int}\Omega_i \neq \phi$, $i = 1, 2, \ldots, n-1$. By Proposition 2.3, $\bigcap_{i=1}^{n-1} \operatorname{int}(\Omega_i) \cap \Omega_n \cap U = \phi$. The result follows from the Theorem of Separation of Convex Sets. ■

- If $\dim X < \infty$ then the extremality and sperarion of convex sets are equivalent without the assumption of nonempty interiors (see Theorem 2.8 below).

Corollary 5 *(2.4 extremality critertion for convex sets)*
Let $\Omega_1, \Omega_2, \ldots, \Omega_n$ be nonempty convex subsets of a space X, with $\operatorname{int}(\Omega_i) \neq \phi$, $i = 1, 2, \ldots, n-1$ and $\overline{x} \in \bigcap_{i=1}^{n} \Omega_i$. Then (1) with $U = X$ is equivalent to the extremality of the system $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$.

1.2 (2.1.2) Versions of the Extremal Principle and Supporting Properties

Definition 6 *(2.5 versions of the extremal principle)*
Let $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ be an extremal system in X. We say that:

(i) *$\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ satisfies the ε-extremal principle if for every $\varepsilon > 0$ there exist $x_i \in \Omega_i \cap \overline{x} + \varepsilon\mathbf{B}$ and $x_i^* \in X^*$ such that*

$$x_i^* \in \widehat{N}_\varepsilon(x_i; \Omega_i), \; i = 1, 2, \ldots, n, \tag{2}$$

$$\sum_{i=1}^{n} x_i^* = 0, \qquad \sum_{i=1}^{n} \|x_i^*\| = 1. \tag{3}$$

(ii) *$\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ satisfies the approximate extremal principle if for every $\varepsilon > 0$ there exist $x_i \in \Omega_i \cap \overline{x} + \varepsilon\mathbf{B}$ and*

$$x_i^* \in \widehat{N}(x_i; \Omega_i) + \varepsilon\mathbf{B}^*, \; i = 1, 2, \ldots, n$$

such that (3) holds.

(iii) $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ *satisfies the exact extremal principle if there exist*

$$x_i^* \in N(\overline{x}; \Omega_i), \;\; i = 1, 2, \ldots, n$$

such that (3) holds.

- Since $\widehat{N}(x;\Omega) + \varepsilon\mathbf{B}^* \subset \widehat{N}_\varepsilon(x;\Omega)$, the ε-extremal principle follows from the approximate extremal principle.
- The relations of the extremal principle provide necessary conditions for local extermal points of systems of sets. In this way they can be viewed as a generealized Eulre-Lagrange equations in a geometric setting.
- They can also be viewed as the nonconvex counterparts to local separation principles of convex sets. For example, the exact separation priciple of the two sets Ω_1, Ω_2 means that there is $x^* \in X^*$ such that

  $$0 \neq x^* \in N(\overline{x}; \Omega_1) \cap (-N(\overline{x}; \Omega_2)).$$

  When both Ω_1, Ω_2 are convex, this reduces to

  $$\langle x^*, u_1 \rangle \leq \langle x^*, u_2 \rangle \;\; \forall u_1 \in \Omega_1, u_2 \in \Omega_2,$$

  which is exactly the classical separation property for convex sets.
- In contrast to classical separation, the extremal principle applies locally. This has two implications
  - Any sufficient condition for convex separation implies set exremality.
  - In many situations, the local extremality can be checked directly without having to check interiority, even for convex sets.

Proposition 7 *(2.6 approximate supporting properties of nonconvex sets)*
Suppose Ω is a proper subset of X and $\overline{x} \in \operatorname{bd}(\Omega)$. Then

(i) *If the ε-extremal principle holds for $\{\Omega, \{\overline{x}\}; \overline{x}\}$ then, for every $\varepsilon > 0$ there is an $x \in (\overline{x} + \varepsilon\mathbf{B}) \cap \operatorname{bd}(\Omega)$ such that $\widehat{N}_\varepsilon(x;\Omega) \backslash M\mathbf{B}^* \neq \{0\}$ for all $M > \varepsilon$.*

(ii) *If the approximate extremal principle holds for $\{\Omega, \{\overline{x}\}; \overline{x}\}$ then, for every $\varepsilon > 0$ there is an $x \in (\overline{x} + \varepsilon\mathbf{B}) \cap \operatorname{bd}(\Omega)$ such that $\widehat{N}(x;\Omega) \neq \{0\}$.*

Therefore, for every proper closed subset Ω of X, the validity of the approximate extremal principle (ε-extremal principle) in X implies the density of the set

$$\left\{x \in \operatorname{bd}(\Omega) : \widehat{N}(x;\Omega) \neq \{0\}\right\} \tag{4}$$

and the set

$$\left\{x \in \operatorname{bd}(\Omega) : \widehat{N}_\varepsilon(x;\Omega) \backslash M\mathbf{B}^* \neq \{0\}\right\}$$

for every $\varepsilon > 0$ and every $M > \varepsilon$.

- If Ω is convex, then (4) describes the supporting points of Ω. This is the content of the celebrated Bishop-Phelps theorem.
- The next proposition deals with the reverse question: the possibility to derive extremal principles from dense sets.

Proposition 8 *(2.7 charachterizations of supporting properties)*
Let M be a Banach space, $\varepsilon \geq 0$ and $M \geq \varepsilon$. TFAE

(a) *For every proper closed subset $\Omega \subset X$, there exists $x \in \mathrm{bd}(\Omega)$ satisfying $\widehat{N}_\varepsilon(x;\Omega) \setminus M\mathbf{B}^* \neq \{0\}$ ($\widehat{N}(x;\Omega) \neq \{0\}$ if $\varepsilon = 0$).*

(b) *For every pair Ω_1, Ω_2 of subsets of X such that $\Omega_1 - \Omega_2$ is proper and closed around the origin there are $x_1 \in \Omega_1$ and $x_2 \in \Omega_2$ such that*

$$0 \in \widehat{N}_\varepsilon(x_1;\Omega_1) \setminus M\mathbf{B}^* + \widehat{N}_\varepsilon(x_2;\Omega_2).$$

1.3 (2.1.3) Extremal Principle in Finite Dimensions

Lemma 9 *Let X be finite dimensional (equipped with the Euclidean norm) and, $\Omega \subset X$ be proper and closed , $x \notin \Omega$, $w \in \prod(x;\Omega)$. Then $x - w \in \widehat{N}(w;\Omega)$.*

Proof. We have, for any u in Ω,

$$\|x-w\|^2 \leq \|x-u\|^2.$$

Then

$$\begin{aligned} 2\langle x-w, u-w\rangle &\leq \|u\|^2 - \|w\|^2 = \langle u+w, u-w\rangle \\ &= 2\langle w, u-w\rangle + \|u-w\|^2. \end{aligned}$$

Hence

$$\langle x-w, u-w\rangle \leq \frac{1}{2}\|u-w\|^2.$$

So,

$$\left\langle x-w, \frac{u-w}{\|u-w\|}\right\rangle \leq \frac{1}{2}\|u-w\|.$$

Thus

$$\overline{\lim_{u \xrightarrow{\Omega} w}} \left\langle x-w, \frac{u-w}{\|u-w\|}\right\rangle \leq 0.$$

■

Theorem 10 *(2.8 extermal principle in finite dimensions)*
The exact extremal principle holds in any space X such that $\dim X < \infty$.

Proof. Assume $\{\Omega_1, \Omega_2, \ldots, \Omega_n; \overline{x}\}$ is an extremal system in X (where we assume wlog that $U = X$) with $\Omega_1, \Omega_2, \ldots, \Omega_n \subset X$ closed and proper. Let $\Omega = \Omega_1 \times \Omega_2 \ldots \times \Omega_n$ and, for $x \in X$, put $x_D = (x, x, \ldots, x)$. Choose $a = (a_1, a_2, \ldots, a_n)$ such that

$$x_D + a \notin \Omega \ \forall x \in X. \tag{5}$$

Define $\rho : X \times \Omega \to \mathbb{R}$ by

$$\rho(x, u) := \|x_D - u + a\| + \|x - \overline{x}\|^2,$$

where, for $v = (v_1, v_2, \ldots, v_n) \in X^n$ we take $\|v\| = \sqrt{\|v_1\|^2 + \|v_2\|^2 + \ldots + \|v_n\|^2}$. Observe that, because of (5) $\|x_D - u + a\| > 0$ for all $u \in \Omega$ and all $x \in X$. Furthermore, for $w \in \prod(x_D + a; \Omega)$, $\rho(x, w)$ has a minimizer $(\widehat{x}, \widehat{w})$ since it is continuous and has bounded level sets. Then, for any $u \in \Omega$,

$$\rho(x, u) \geq \rho(x, w) \geq \rho(\widehat{x}, \widehat{w}).$$

In particular

$$\rho(x, \widehat{w}) \geq \rho(\widehat{x}, \widehat{w})$$

and $\widehat{x}$ is a minmizer of $\rho(x, \widehat{w})$. Since $\rho(x, \widehat{w})$ is differentiable, we must have

$$\frac{\partial \rho}{\partial x}\Big|_{(\widehat{x}, \widehat{w})} = 0.$$

This last condition gives

$$\sum_{i=1}^{n} \frac{\widehat{x} + a_i - \widehat{w}_i}{\|x_D - u + a\|} + 2(\widehat{x} - \overline{x}) = 0,$$

which will be written in short as

$$\sum_{i=1}^{n} x_i^*(a) + 2(\widehat{x} - \overline{x}) = 0 \tag{6}$$

with

$$x_i^*(a) = \frac{\widehat{x} + a_i - \widehat{w}_i}{\|\widehat{x}_D - \widehat{w} + a\|}.$$

Putting

$$x^*(a) = \frac{\widehat{x}_D - \widehat{w} + a}{\|\widehat{x}_D - \widehat{w} + a\|} = (x_1^*(a), x_2^*(a), \ldots, x_n^*(a)),$$

we get from Lemma 9 that $x^*(a) \in \widehat{N}(\widehat{w}; \Omega)$. It follows that for $i = 1, 2, \ldots, n$, $x_i^*(a) \in \widehat{N}(\widehat{w}_i; \Omega_i)$. Also it is easy to check that

$$\|x^*(a)\|^2 = \sum_{i=1}^{n} \|x_i^*(a)\|^2 = 1. \tag{7}$$

Now take a sequence $a_k \to 0$ such that (5) is satisfied for each k. We get that (a subsequence) $(\widehat{x}_k, \widehat{w}_k) \to (\overline{x}, \overline{x})$. Since the unit sphere in $X^{*n} = X^n$ is compact, (a subsequence) $x^*(a_k) \to x^* = (x_1^*, x_2^*, \ldots, x_n^*)$. Thus, $x^* \in N(\overline{x}_D; \Omega)$ and $x_i^* \in N(\overline{x}; \Omega_i)$, $i = 1, 2, \ldots, n$. Taking the limit in equation (6) gives

$$\sum_{i=1}^{n} x_i^* = 0.$$

Taking the limit in equation (7) gives

$$\|x^*\|^2 = \sum_{i=1}^{n} \|x_i^*\|^2 = 1.$$

■