INTRODUCTION TO VARIATIONAL ANALYSIS

BORIS MORDUKHOVICH

Department of Mathematics

Wayne State University

www.math.wayne.edu/∼boris

LECTURES presented at the

CIMPA-UNESCO SCHOOL on

Variational Inequalities and Related Problems

Hanoi, Vietnam

May 2010

THE FUNDAMENTAL VARIATIONAL PRINCIPLE

...Namely, because the shape of the whole universe is the most perfect and, in fact, designed by the wisest creator, nothing in all the world will occur in which no maximum or minimum rule is somehow shining forth...

Leonhard Euler (1744)

NONSMOOTHNESS IN OPTIMIZATION AND EQUILIBRIA

–**max-functions**

$$f(x) = \max_{u \in U} g(x,u)$$

—**distance functions**

$$d(x;\Omega) = \inf_{u \in \Omega} \|x-u\|, \;\; \rho(x,y) = \inf_{u \in F(y)} \|x-u\|$$

–**value functions**

$$V(\alpha) = \inf\{f(x,\alpha)|\; h(x,\alpha)=0, \;\; g(x,\alpha) \le 0\}$$

–sets of *feasible and optimal solutions* in **parametric optimization**

–*feasible and optimal allocations* in models of **economic equilibria**

–*reachable sets* in **optimal control**.

CONVEX FUNCTIONS

$f\colon I\!\!R^n \to \overline{I\!\!R} := (-\infty, \infty]$ is **convex** if its *epigraphical* set

$$\mathrm{epi}\, f = \{(x, \mu) \in I\!\!R^n \times I\!\!R |\ \mu \geq f(x)\}$$

is *convex*. **Subdifferential** of f at $\bar{x}$ is

$$\partial f(\bar{x}) = \{x^* |\ \langle x^*, x - \bar{x}\rangle \leq f(x) - f(\bar{x})\ \forall x \in I\!\!R^n\}$$

Directional derivative of f at $\bar{x}$

$$f'(\bar{x}; v) = \lim_{t \downarrow 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t}$$

Dual representation

$$\partial f(\bar{x}) = \{x^* |\ \langle x^*, x\rangle \leq f'(\bar{x}; v)\ \ \forall v \in I\!\!R^n\}$$

Sum Rule

$$\partial(f_1 + f_2)(\bar{x}) = \partial f_1(\bar{x}) + \partial f_2(\bar{x})$$

provided that one of f_i is *continuous* around $\bar{x}$.

CONVEX GEOMETRY

Normal cone to **convex** $\Omega \subset I\!R^n$ at $\bar{x} \in \Omega$

$N(\bar{x};\Omega) = \{x^* \in I\!R^n |\ \langle x^*, x-\bar{x}\rangle \le 0 \ \ \forall x \in \Omega\}$

or, via the *indicator function*,

$$N(\bar{x};\Omega) = \partial\delta(\bar{x};\Omega)$$

Both the *subdifferential and normal cone* are closed and *convex* sets.

Intersection Rule

$$N(\bar{x};\Omega_1 \cap \Omega_2) = N(\bar{x};\Omega_1) + N(\bar{x};\Omega_2)$$

provided that $\Omega_1 \cap \operatorname{int}\Omega_2 \neq \emptyset$

Everything in convex analysis is based on separation theorems for convex sets.

GENERALIZED GRADIENT

Clarke directional derivative (1975) for a locally *Lipschitz* function $f\colon \mathbb{R}^n \to \mathbb{R}$

$$f^\circ(\bar{x}; v) = \limsup_{\substack{x \to \bar{x} \\ t \downarrow 0}} \frac{f(x+tv) - f(\bar{x})}{t}$$

Generalized gradient of f at $\bar{x}$

$$\partial_C f(\bar{x}) = \{x^* |\ \langle x^*, v\rangle \le f^\circ(\bar{x}; v) \ \ \forall v \in \mathbb{R}^n\}$$

Both $f^\circ(\bar{x}, \cdot)$ and $\partial_C f(\bar{x})$ are **convex**.

Sum Rule for Lipschitz functions

$$\partial_C(f_1 + f_2)(\bar{x}) \subset \partial_C f_1(\bar{x}) + \partial_C f_2(\bar{x})$$

All proofs are based on separation.

Two-sided construction: **symmetry**

$$\partial_C(-f)(\bar{x}) = -\partial_C f(\bar{x})$$

RELATED GEOMETRY

Clarke tangent cone to $\Omega \subset I\!R^n$ at $\bar{x} \in \Omega$

$$T_c(\bar{x};\Omega) = \{v|\ \forall x_k \to \bar{x}, \forall t_k \downarrow 0\ \exists v_k \to v \quad \text{with}\ \ x_k + t_k v_k \in \Omega\}$$

Clarke normal cone

$$N_c(\bar{x};\Omega) = \{x^*|\ \langle x^*, v\rangle \le 0\ \forall v \in T_c(\bar{x};\Omega)\}$$

Both $T_c(\bar{x})$ and $N_c(\bar{x};\Omega)$ are **convex**.

Theorem (Rockafellar, 1985). *Let Ω be the* **graph** *of a locally* **Lipschitz** *function $f\colon I\!R^n \to I\!R^m$. Then $N_c((\bar{x}, f(\bar{x}));\Omega)$ is a* **linear subspace** *of dimension $q \ge m$ with $q = m$* **if and only if** *f is* **smooth** *around $\bar{x}$. Holds also when Ω is a* **Lipschitz manifold** (*e.g., graphs of* **maximal monotone** *maps.*)

DISCUSSIONS ON CONVEXITY

Clarke's generalized gradient is the **smallest** among any **convex-valued** subdifferentials for *Lipschitz functions* with reasonable properties including **robustness** (closed graph)

$$\partial_C f(\bar{x}) = \operatorname*{Lim\,sup}_{x \to \bar{x}} \partial_C f(x)$$

where "lim sup" stands for the *Painlevé-Kuratowski* **upper/outer limit** of $F\colon X \rightrightarrows Y$ as $x \to \bar{x}$

$$\operatorname*{Lim\,sup}_{x \to \bar{x}} F(x) = \{y \mid \exists x_k \to \bar{x}, \exists y_k \to y \quad \text{with} \;\; y_k \in F(x_k)\}$$

But $\partial_C f$ has **serious drawbacks** in variational analysis: **too large** for *optimality conditions*, *weak chain rules*, *fails on corners* (as *complementarity*), *doesn't apply to MPECs*...

ABSTRACT SUBDIFFERENTIALS

of *extended-real-valued* functions $f : R^n \to \overline{I\!R}$

Natural Requirements

1) for **convex** f

$$\partial^\bullet f(\bar{x}) = \{x^* |\ f(x) - f(\bar{x}) \le \langle x^*, x - \bar{x}\rangle \forall x \in R^n\}$$

2) If $\bar{x}$ is a **local minimizer**, then

$$0 \in \partial^\bullet f(\bar{x})$$

3) **Sum Rule**

$$\partial^\bullet (f_1 + f_2)(\bar{x}) \subset \partial^\bullet f_1(\bar{x}) + \partial^\bullet f_2(\bar{x})$$

4) **Robustness**

$$\partial^\bullet f(\bar{x}) = \operatorname*{Lim\,sup}_{x \to \bar{x}} \partial^\bullet f(x)$$

BASIC/LIMITING SUBDIFFERENTIAL

of *l.s.c.* functions is defined by (Mor1976)

$$\partial f(\bar{x}) = \operatorname*{Lim\,sup}_{x \xrightarrow{f} \bar{x}} \widehat{\partial} f(x)$$

where $x \xrightarrow{f} \bar{x} \Longleftrightarrow x \to \bar{x}, f(x) \to f(\bar{x})$ and

$$\widehat{\partial} f(x) = \{v | \liminf_{u \to x} \frac{f(u) - f(x) - \langle v, u - x \rangle}{\|u - x\|} \geq 0\}$$

It is **minimal** among all subdifferentials satisfying 1)-4), often **nonconvex**, and

$$\partial f(\bar{x}) \neq \emptyset \text{ for Lipschitz functions}$$

where the *convexification* of $\partial f(\bar{x})$ gives *Clarke's generalized gradient* $\partial_C f(\bar{x})$.

Example: $\partial(-|x|)(0) = \{-1, 1\}$ while

$$\partial_C(-|x|)(0) = \partial_C(|x|)(0) = [-1, 1]$$

FULL CALCULUS

Besides **Sum Rule**, we have **Chain Rule** for $f \circ g$ with $f\colon I\!R^m \to I\!R$ and $g\colon I\!R^n \to I\!R^m$

$$\partial(f \circ g)(\bar{x}) \subset \bigcup_{y^* \in \partial f(\bar{y})} \partial\langle y^*, g\rangle(\bar{x}), \quad \bar{y} = g(\bar{x})$$

Product Rule with $\lambda_i = f_i(\bar{x})$

$$\partial(f_1 \cdot f_2)(\bar{x}) \subset \partial(\lambda_2 f_1)(\bar{x}) + \partial(\lambda_1 f_2)(\bar{x})$$

Quotient Rule, Maximum Rule, Minimum Rule, Partial Differentiation, Mean Value Theorem

$$f(b) - f(a) = \langle x^*, b - a\rangle$$

with $x^* \in \partial f(c) \cup (-\partial(-f)(c))$, $c \in (a, b)$

All proofs are based on the **Extremal Principle** which is a **variational analog** of local **separation without convexity**; see below.

VARIATIONAL GEOMETRY

The (basic, limiting, M- **normal cone** $N(\bar{x};\Omega) = \partial\delta(\bar{x};\Omega)$ can be *equivalently* defined by

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x;\Omega)$$

where the **prenormal cone** (of *Fréchet or regular normals*) is

$$\widehat{N}(x;\Omega) = \{x^* |\ \limsup_{u \xrightarrow{\Omega} x} \frac{\langle x^*, u - x\rangle}{\|x - u\|} \leq 0\}$$

Equivalent definition of limiting normals (Mor76)

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x \rightarrow \bar{x}} [\text{cone}\,(x - \Pi(x;\Omega))]$$

where $\Pi(x;\Omega)$ is the *Euclidean projector.* Then

$$\partial f(\bar{x}) = \{x^* |\ (x^*, -1) \in N((\bar{x}, f(\bar{x})); \text{epi}\, f)\}$$

Clarke's normal cone is

$$N_c(\bar{x};\Omega) = \text{clco}\, N(\bar{x};\Omega)$$

RELATIONS WITH TANGENTS

The basic **normal cone** is **not convex**, hence **cannot be tangentially generated**. But for the **prenormal cone** we have

$$\widehat{N}(\bar{x};\Omega) = \{x^*|\ \langle x^*, v\rangle \le 0\ \forall v \in T(\bar{x};\Omega)\}$$

where $T(\bar{x};\Omega)$ is the (Bouligand-Severi, 1930) **contingent cone**

$$T(\bar{x};\Omega) = \{v|\ \exists x_k \to \bar{x},\ \exists \alpha_k \to \infty \\ \text{with}\ \ \alpha_k(x_k - \bar{x}) \to v\}$$

Contingent cone $T(\bar{x};\Omega)$ is often *nonconvex* with

$$T_c(\bar{x};\Omega) \subset T(\bar{x};\Omega)$$

In general $T(\cdot;\Omega)$ and $\widehat{N}(\cdot;\Omega)$ have **poor calculus**; for $T_c(\cdot)$ and $N_c(\cdot;\Omega)$ is **better but not satisfactory**; $N(\cdot;\Omega)$ has **all we need**.

BASIC CALCULUS RULE

Theorem. *Let* Ω_1 *and* Ω_2 *be closed around* $\bar{x} \in \Omega_1 \cap \Omega_2$. *Then*

$$N(\bar{x};\Omega_1 \cap \Omega_2) \subset N(\bar{x};\Omega_1) + N(\bar{x};\Omega)$$

provided the **qualification condition**

$$N(\bar{x};\Omega_1) \cap (-N(\bar{x};\Omega_2)) = \{0\}$$

Moreover

$$N(\bar{x};\Omega_1 \cap \Omega_2) = N(\bar{x};\Omega_1) + N(\bar{x};\Omega_2)$$

if both Ω_i *are* **normally regular** *at* $\bar{x}$, *i.e.*,

$$N(\bar{x};\Omega_i) = \widehat{N}(\bar{x};\Omega_i), \quad i = 1, 2$$

Proof is based on the Extremal Principle; see below.

SET EXTREMALITY

Definition. $\bar{x} \in \Omega_1 \cap \Omega_2$ is a **local extremal point** of the set system $\{\Omega_1, \Omega_2\}$ if there exists a neighborhood U of $\bar{x}$ such that for any $\varepsilon > 0$ there is $a \in X$ with $\|a\| < \varepsilon$ satisfying

$$(\Omega_1 + a) \cap \Omega_2 \cap U = \emptyset$$

Examples:

–*boundary point* of closed sets

–*local solutions* to *constrained optimization* and *multiobjective optimization* problems

–*minimax solutions* and *equilibrium points*

–*Pareto optimal allocations in economics*

–*stationary points in mechanical models*, etc.

EXTREMAL POINTS IN OPTIMIZATION

Constrained Optimization

$$\text{minimize} \ \ f(x) \ \ \text{subject to} \ \ x \in \Omega \subset I\!R^n$$

$\bar{x}$ is a *local minimizer*. Then the point $(\bar{x}, f(\bar{x}))$ is *locally extremal* for $\{\Omega_1, \Omega_2\}$ in $I\!R^n \times I\!R$ with the sets Ω_i defined by

$$\Omega_1 = \operatorname{epi} f = \{(x, \mu) |\ \mu \ge f(x)\}$$

$$\Omega_2 = \Omega \times \{f(\bar{x})\}$$

Indeed, given $\varepsilon > 0$ take $a = (0, -\varepsilon) \in I\!R^{n+1}$.

—For closed **convex** sets the notions of **extremality and separation are equivalent**

—In **optimization** and **equilibrium** problems (*not only*) **extremality is automatic**.

EXTREMAL PRINCIPLE

Theorem (Mor1976, Kruger-Mor1980). *Let* $\bar{x}$ *be a* **locally extremal point** *for the system of closed sets* $\{\Omega_1, \Omega_2\}$ *in* R^n. *Then there exists* $x^* \neq 0$ *such that*

$$x^* \in N(\bar{x}; \Omega_1) \cap (-N(\bar{x}; \Omega_2))$$

which is a **generalized Euler equation**.

This is a **variational counterpart** of the **separation theorem** in the case of **nonconvex** sets; goes back to *classical separation* when both Ω_i are *convex* due to the structure of the normal cone

$$N(\bar{x}; \Omega) = \{x^* | \ \langle x^*, x - \bar{x} \rangle \leq 0 \ \forall x \in \Omega\}$$

PROOF OF THE EXT. PRINCIPLE (METRIC APPROXIMATIONS)

Let $a_k \to 0$. Consider a family of **unconstrained minimization** problems (P_k):

$$\rho_k(x) = [d^2(x + a_k; \Omega_1) + d^2(x; \Omega_2)]^{1/2} + \|x - \bar{x}\|^2$$

By Weierstrass $\exists$ *minimizers* x_k to (P_k) with

$$\alpha_k := [d^2(x_k + a_k; \Omega_1) + d^2(x_k; \Omega_2)]^{1/2} \to 0$$

and $x_k \to \bar{x}$. By **extremality** $\alpha_k > 0$ (!) Take $w_{1k} \in \Pi(x_k + a_k; \Omega_1)$, $w_{2k} \in \Pi(x_k; \Omega_2)$ and form the function

$$\begin{aligned} r_k(x) = \ & [\|x + a_k - w_{1k}\|^2 + \|x - w_{2k}\|^2]^{1/2} \\ & + \|x - \bar{x}\|^2 \end{aligned}$$

which is **smooth** around x_k.

PROOF OF THE EXT. PRINCIPLE (CONTINUATION)

Clearly x_k **minimizes** $r_k(\cdot)$. By **Fermat**

$$\nabla r_k(x_k) = x^*_{1k} + x^*_{2k} + 2(x_k - \bar{x}) = 0$$

where

$$x^*_{1k} = \frac{x_k + a_k - w_{1k}}{\alpha_k}, \quad x^*_{2k} = \frac{x_k - w_{2k}}{\alpha_k}$$

Then

$$\|x^*_{1k}\|^2 + \|x^*_{2k}\|^2 = 1 \text{ for all } k$$

and there are $x^*_1, x^*_2 \in I\!R^n$ such that

$$x^*_{ik} \to x^*_i \text{ for } i = 1, 2$$

Passing to the limit as $k \to \infty$ and taking into account the construction of the normal cone as

PROOF OF THE EXT. PRINCIPLE (COMPLETION)

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x \to \bar{x}} \left[\operatorname{cone}\left(x - \Pi(x;\Omega)\right)\right]$$

we get

$$x_i^* \in N(\bar{x};\Omega_i), \quad i = 1,2$$

$$x_1^* + x_2^* = 0, \quad \|x_1^*\|^2 + \|x_2^*\|^2 = 1$$

The proof is completed by putting

$$x^* := x_1^* = -x_2^*$$

Corollary. *Let* $\Omega \subset I\!\!R^n$ *be closed. Then*

$$N(\bar{x};\Omega) \neq \{0\}$$

if and only if $\bar{x}$ *is a* **boundary point** *of the set* Ω.

PROOF OF SUM RULE (SKETCH)

Take $x^* \in \partial(f_1 + f_2)(\bar{x})$ and for any $\varepsilon_k \downarrow 0$, find $x_k \to \bar{x}$ and $x_k^* \to x^*$ such that

$$(f_1 + f_2)(x) - (f_1 + f_2)(x_k) - \langle x^*, x - x_k \rangle$$
$$> -\varepsilon_k \|x - x_k\| \ \text{ near } \ x_k \ \text{ for each } \ k$$

Form two closed sets

$$\Omega_{1k} = \{(x, \mu)|\ f_1(x) - f_1(x_k) \le \mu\}$$

$$\Omega_{2k} = \{(x, \mu)|\ f_2(x) - f_2(x_k) - \langle x_k^*, x - x_k \rangle + \varepsilon_k \|x - x_k\| \le -\mu\}$$

and observe that $(x_k, 0)$ is a **locally extremal point** of $\{\Omega_{1k}, \Omega_{2k}\}$ for each k. Employing the **extremal principle** and *passing to the limit* as $k \to \infty$ (invoking the **robustness** property of $N(\cdot;\Omega)$), we complete the proof.

CODERIVATIVES

Let $F : X \rightrightarrows Y$ (*set-valued*) with $\bar{y} \in F(\bar{x})$. Then $D^*F(\bar{x},\bar{y}) : Y^* \rightrightarrows X^*$ defined in [Mor80]

$$D^*F(\bar{x},\bar{y})(y^*) = \{x^* | (x^*,-y^*) \in N(\bar{x},\bar{y};\operatorname{gph} F)\}$$

is called the **coderivative** of F at $(\bar{x},\bar{y})$.

If $F : X \to Y$ is *smooth* (strictly differentiable, C^1) at $\bar{x}$, then

$$D^*F(\bar{x})(y^*) = \{(\nabla F(\bar{x}))^* y^*\} \ \forall y^* \in Y^*$$

i.e., the coderivative is a proper generalization of the classical **adjoint derivative** operator. If $F : X \to Y$ is *locally Lipschitzian* at $\bar{x}$, then the **scalarization formula** holds

$$D^*F(\bar{x})(y^*) = \partial\langle y^*, F\rangle(\bar{x})$$

Full Calculus!–Sum, Chain Rules, etc.

METRIC REGULARITY

A set-valued mapping $F(\cdot)$ is **metrically regular** around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$, and positive numbers μ and ε such that

$$d(x; F^{-1}(y)) \leq \mu\, d(y; F(x))$$

for all $x \in U$ and $y \in V$ with $d(y; F(x)) \leq \varepsilon$. The **exact regularity bound** is

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \inf\{\mu\}$$

Goes back to Banach (1931, *linear* operators), Lyusternik (1934) and Graves (1950) for *smooth nonlinear* mappings. *Equivalent* to **covering** or **openness at linear rate** with

$$\operatorname{cov} F(\bar{x}, \bar{y}) = 1/\operatorname{reg} F(\bar{x}, \bar{y})$$

LIPSCHITZIAN PROPERTIES

A set-valued mapping $F\colon I\!R^n \rightrightarrows I\!R^m$ is called **Lipschitz-like** (or having the *Aubin property*) around $(\bar{x},\bar{y}) \in \operatorname{gph} F$ with modulus ℓ if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$F(x) \cap V \subset F(u) + \ell\|x-u\| I\!B$$

for all $x, u \in U$, where $I\!B$ is the closed unit ball. *Equivalent* to local *Lipschitz continuity* of

$$\rho(x,y) = d(y; F(x)).$$

The **exact Lipschitzian bound** of F around $(\bar{x},\bar{y})$ is

$$\operatorname{lip} F(\bar{x},\bar{y}) = \inf\{\ell\}$$

Reduces to the *classical Lipschitzian* property of F around $\bar{x}$ for $V = I\!R^m$ and *compact* $F(x)$.

EQUIVALENCY RELATIONS

Theorem. *A set-valued mapping F is* **Lipschitz-like** *around* $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ **if and only if** *the* **inverse mapping**

$$F^{-1}(y) := \{x|\ y \in F(x)\}$$

is **metrically regular** *around* $(\bar{y}, \bar{x}) \in \operatorname{gph} F^{-1}$. *In this case*

$$\operatorname{lip} F(\bar{x}, \bar{y}) = \operatorname{reg} F^{-1}(\bar{y}, \bar{x})$$

There are **semi-local** (with respect to either $\bar{x}$ or $\bar{y}$) versions of these notions and results.

CODERIVATIVE CRITERIA

Theorem (Mor1984,Mor1993). *A set-valued mapping* $F\colon I\!R^n \rightrightarrows I\!R^m$ *is* **Lipschitz-like** *around* $(\bar{x},\bar{y})$ **if and only if**

$$D^*F(\bar{x},\bar{y})(0)=\{0\}$$

In this case

$\operatorname{lip} F(\bar{x},\bar{y})=\|D^*F(\bar{x},\bar{y})\|$

$=\sup\{\|x^*\| \;\; s.t. \;\; x^*\in D^*F(\bar{x},\bar{y})(y^*),\; \|y^*\|=1\}$

Equivalently, F *is* **metrically regular** *around* $(\bar{x},\bar{y})$ **if and only if**

$$\ker D^*F(\bar{x},\bar{y})=\{0\}$$

which reduces to the **surjectivity condition** $\nabla F(\bar{x})I\!R^n=I\!R^m$ *in the smooth case.*

Great many applications!

NONDIFFERENTIABLE PROGRAMMING

Minimize $\varphi_0(x)$ subject to $x \in \Omega$ and

$$\varphi_i(x) \leq 0,\ i = 1, \ldots, m$$

$$\varphi_i(x) = 0,\ i = m+1, \ldots, m+r$$

Theorem. *Let* φ_i *be* **Lipschitzian** *and* Ω *closed around an optimal solution* $\bar{x}$. *Then there are* $(\lambda_0, \ldots, \lambda_{m+r}) \neq 0$ *satisfying the conditions* $\lambda_i \geq 0,\ i = 0, \ldots, m,$

$$\lambda_i \varphi_i(\bar{x}) = 0,\ i = 1, \ldots, m$$

$$0 \in \partial(\Sigma_{i=0}^{m+r} \lambda_i \varphi_i)(\bar{x}) + N(\bar{x}; \Omega)$$

Moreover, $\lambda_0 = 1$ *under nonsmooth extensions of the* **Mangasarian-Fromovitz Constraint Qualifications (MFCQ)**.

MATH PROGRAMS WITH EQUILIBRIUM CONSTRAINTS

Luo-Pang-Ralph (1996), Outrata-Kočvara-Zowe (1998), Facchinei-Pang (2003)

General MPEC Model

$$\text{minimize} \;\; \varphi(x,y) \;\; \text{subject to}$$

$$0 \in f(x,y) + Q(x,y)$$

under **generalized equation (GEs)** constraints (Robinson, 1979). In particular, **variational inequality** constraints as $Q(y) = N(y;\Omega)$ with a *convex* set Ω are: *find*

$$y \in \Omega \;\; with \;\; \langle f(x,y), u - y \rangle \geq 0 \;\; \forall u \in \Omega$$

Complementarity constraints correspond to $\Omega = I\!\!R^n_+$. Covers **bilevel programming**.

NECESSARY OPTIMALITY CONDITIONS FOR MPECs

Theorem. *Let* $(\bar{x}, \bar{y})$ *be an optimal solution to the general MPEC:*

$$\text{minimize } \varphi(x,y) \text{ subject to}$$
$$0 \in f(x,y) + Q(x,y)$$

where φ *and* f *are* Lipschitz *around* $(\bar{x}, \bar{y})$*. Let the* **adjoint generalized equation**

$$0 \in \partial\langle z^*, f\rangle(\bar{x},\bar{y}) + D^*Q(\bar{x},\bar{y},\bar{z})(z^*)$$

with $\bar{z} = -f(\bar{x},\bar{y})$ *admit* **only the trivial solution** $z^* = 0$*. Then there are vectors* $(x^*, y^*) \in \partial\varphi(\bar{x},\bar{y})$ *and* z^* *satisfying*

$$(-x^*, -y^*) \in \partial\langle z^*, f\rangle(\bar{x},\bar{y}) + D^*Q(\bar{x},\bar{y},\bar{z})(z^*)$$

2ND-ORDER SUBDIFFERENTIALS

Let $\varphi: I\!R^n \to \overline{I\!R}$ at $\bar{x}$ and $\bar{y} \in \partial\varphi(\bar{x})$. Then

$$\partial^2\varphi(\bar{x},\bar{y})(y) := (D^*\partial\varphi)(\bar{x},\bar{y})(u)$$

is the **second-order subdifferential** of φ at $\bar{x}$ relative to $\bar{y}$ (Mor92). If $\varphi \in C^2$ near $\bar{x}$, then

$$\partial^2\varphi(\bar{x})(u) = \{\nabla^2\varphi(\bar{x})^*u\} = \{\nabla^2\varphi(\bar{x})u\}$$

Full calculus is available

Efficiently computed for various *attractive classes of nonsmooth functions.* In particular

—indicators of **polyhedral convex sets** (Dontchev-Rockafellar, 1996; Yao-Yen, 2009; Henrion-Mor-Nam, 2010)

—separable **piecewise C^2 functions** (Mor-Outrata, 2001, 2007)

COMPOSITE MPECs I

minimize $\varphi(x,y)$ subject to

$$0 \in f(x,y) + \partial(\psi \circ g)(x,y)$$

Theorem. *Let $(\bar{x},\bar{y})$ be an optimal solution, f be $\mathcal{C}^1$ with* **surjective** $\nabla_x f(\bar{x},\bar{y})$, *$g = g(y)$ be $\mathcal{C}^2$ with surjective $\nabla g(\bar{y})$, and $\bar{v}$ be a unique solution to*

$$-f(\bar{x},\bar{y}) = \nabla g(\bar{y})^*\bar{v}, \quad \bar{v} \in \partial\psi(\bar{w}),\ \bar{w} = g(\bar{y})$$

Assume that φ is **Lipschitz** *continuous around $(\bar{x},\bar{y})$. Then there are (x^*,y^*,u) satisfying*

$$(x^*,y^*) \in \partial\varphi(\bar{x},\bar{y}), \quad -x^* = \nabla_x f(\bar{x},\bar{y})^* u,$$

and the **2nd-order generalized equation**

$$-y^* \in \nabla_y f(\bar{x},\bar{y})^* u + \nabla^2\langle\bar{v},g\rangle(\bar{y})^* u$$
$$+\nabla g(\bar{y})^*\partial^2\psi(\bar{w},\bar{v})(\nabla g(\bar{y})u)$$

COMPOSITE MPECs II

minimize $\varphi(x,y)$ subject to

$$0 \in f(x,y) + (\partial\psi \circ g)(x,y)$$

Theorem. *Let $(\bar{x},\bar{y})$ be an optimal solution to the MPEC, let vector functions f and g be $\mathcal{C}^1$, and let the* **second-order qualification condition**

$$\partial^2\psi(\bar{w},\bar{z})(0) \cap \ker \nabla g(\bar{x},\bar{y})^* = \{0\}$$

holds, where $\bar{w} = g(\bar{x},\bar{y})$ and $\bar{z} = -f(\bar{x},\bar{y})$. Then there are vectors $(x^*,y^*) \in \partial\varphi(\bar{x},\bar{y})$ and u satisfying

$$-x^* \in \nabla_x f(\bar{x},\bar{y})^* u + \nabla_x g(\bar{x},\bar{y})^* \partial^2\psi(\bar{w},\bar{z})(u)$$

$$-y^* \in \nabla_y f(\bar{x},\bar{y})^* u + \nabla_y g(\bar{y})^* \partial^2\psi(\bar{w},\bar{z})(u)$$

STABLE CONSTRAINT SYSTEMS

Theorem. *Let*

$$F(x) = \{y|\ \varphi_i(x,y) \le 0,\ i = 1,\ldots,m$$

$$\varphi_i(x,y) = 0,\ i = m+1,\ldots,m+r\}$$

with **smooth** φ_i *under* **MFCQ** *at* $(\bar{x},\bar{y})$, *with* **active constraint** *indices* $I(\bar{x},\bar{y})$. *Then*

$$[\Sigma_{i\in I(\bar{x},\bar{y})}\lambda_i\nabla_y\varphi_i(\bar{x},\bar{y}) = 0]$$

$$\Longrightarrow [\Sigma_{i\in I(\bar{x},\bar{y})}\lambda_i\nabla_x\varphi_i(\bar{x},\bar{y}) = 0]$$

is **necessary and sufficient** *for* F *to be* **Lipschitz-like** *around* $(\bar{x},\bar{y})$. *In this case*

$$\operatorname{lip} F(\bar{x},\bar{y}) = \max\{\|\Sigma_{i\in I(\bar{x},\bar{y})}\lambda_i\nabla_x\varphi_i(\bar{x},\bar{y})\|$$

$$\text{s.t. } \|\Sigma_{i\in I(\bar{x},\bar{y})}\lambda_i\nabla_y\varphi_i(\bar{x},\bar{y})\| \le 1\}$$

Can be generalized to

$$F(x) = \{y|\ g(x,y) \in \Theta,\ \ (x,y) \in \Omega\}$$

SENSITIVITY ANALYSIS FOR VARIATIONAL SYSTEMS

Canonical Perturbations

$$S(x,p) = \{y|\, p \in f(x,y) + Q(x,y)\}$$

Theorem. *Let* $\bar{y} \in S(\bar{x},\bar{p})$, *$f$ be smooth around* $(\bar{x},\bar{y})$, $Q = Q(y)$. *Then S is* **Lipschitz-like** *around* $(\bar{x},\bar{y},\bar{q})$ *with* $\bar{q} = \bar{p} - f(\bar{x},\bar{y})$ **if and only if** *the* **partial adjoint generalized equation**

$$0 \in \nabla_y f(\bar{x},\bar{y})^* z^* + D^*Q(\bar{y},\bar{s})(z^*)$$

has **only the trivial solution** $z^* = 0$.

Extensions and applications to **nonsmooth GEs**, **variational and hemivariational inequalities, complementarity**, etc.

VARIATIONAL PRINCIPLES IN INFINITE DIMENSIONS

Theorem (Ekeland, 1972). *Let* (X, d) *be a* **complete metric** *space,* $\varphi: X \to \overline{I\!R}$ **l.s.c. bounded below**, $\varepsilon > 0$, $x_0 \in X$ *satisfy*

$$\varphi(x_0) \le \inf_X + \varepsilon$$

Then for every $\lambda > 0$ *there is* $\bar{x} \in X$ *with*

$$\varphi(\bar{x}) \le \varphi(x_0)$$

$$\mathrm{dist}(\bar{x}, x_0) \le \lambda$$

$$\varphi(x) + \frac{\varepsilon}{\lambda} d(x, \bar{x}) > \varphi(\bar{x}) \ \forall x \ne \bar{x}$$

Moreover, the **completeness is necessary** *for Ekeland's variational principle.*

Also **smooth variational principles** in **smooth Banach** spaces (Borwein-Preiss, Deville et al.)

ASPLUND SPACES

Originally defined by Asplund (1968) as those Banach spaces on which *every convex continuous function is densely Fréchet differentiable.*

Characterizations: *every separable subspace has separable dual*; every *maximal monotone* operation is *generically single-valued*, X^* has the *Radon-Nicodým* property, etc.

Examples: Banach spaces those *dual are separable*, those with *Fréchet differentiable renorms* or bump functions, *all reflexive spaces*, etc. There are Asplund spaces which do *not* admit even a *Gâteaux smooth renorm.*

APPROXIMATE EXTREMAL PRINCIPLE

Theorem (Mor-Shao1996). *The following are* **equivalent***:*

(i) X *is* **Asplund**

(ii) *For every* **locally extremal point** $\bar{x} \in \Omega_1 \cap \Omega_2$ *of closed sets in* X *and every* $\varepsilon > 0$ *there are are* $x_i \in \Omega_i \cap (\bar{x} + \varepsilon I\!B)$ *and*

$$x_i^* \in \widehat{N}(x_i; \Omega_i) + \varepsilon I\!B^*, \quad i = 1, 2$$

satisfying the **generalized Euler equation**

$$x_1^* + x_2^* = 0, \quad \|x_1^*\| + \|x_2^*\| = 1$$

Extended to **finitely many** sets. The **Exact Extremal Principle** holds under **sequential normal compactness** of all but one sets.

REFERENCES (BOOKS)

1. H. Attouch, G. Buttazzo and G. Michaile, *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*, SIAM, Philadelphia, 2005.

2. J.M. Borwein and Q.J. Zhu, *Techniques in Variational Analysis.* Springer, New York, 2005.

3. F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

4. F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, New York, 2003.

5. V. Jeyakumar and D.T. Luc, *Nonsmooth*

Vector Functions and Continuous Optimization, Springer, New York, 2008.

6. B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Springer, Berlin, 2006.

7. B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation, II: Applications*, Springer, Berlin, 2006.

8. W. Schirotzek, *Nonsmooth Analysis*, Springer, Berlin, 2007.

9. R.T. Rockafellar and R. J-B Wets, *Variational Analysis*, Springer, Berlin, 1998.

10. R.B. Vinter, *Optimal Control*, Birkhäuser, Basel, 2000.