An introduction to Nonsmooth Analysis

Abderrahim Jourani

Université de Bourgogne
Institut de Mathématiques de Bourgogne
UMR 5584 CNRS
B.P. 47870, 21078 Dijon, France

CMM-Universidad de Chile, Santiago, December 2023

Outline of the course

- Motivations and Examples
- Elements of Nonsmooth Analysis
- Density theorem
- Applications
  - In classical analysis : Mean value theorem, Monotonicity theorem.
  - In Optimization : Necessary optimality conditions, Calmness as a constraint qualification condition, Value function,
  - In optimal control: Value function in optimal control, Verification function, Minimal time problem, Necessary optimality conditions of free time problems, Invariance, Hamilton-Jacobi equations

Motivations

Nonsmooth problems arise in many fields of applications, for example in

- *image denoising,*
- *optimal control,*
- *neural network training,*
- *data mining,*
- *economics, and*
- *computational chemistry and physics.*

Moreover, using certain important methodologies for solving difficult smooth problems leads directly to the need to solve nonsmooth problems. This is the case, for instance in

- *decompositions,*
- *dual formulations, and*
- *exact penalty functions.*

Difficulties caused by nonsmoothness

SMOOTH PROBLEM:

• *Descent direction is obtained at the opposite direction gradient $\nabla f(x)$.*

• *The necessary optimality condition $\nabla f(x) = 0$.*

• *Difference approximation can be used to approximate the gradient.*

NONSMOOTH PROBLEM:

• *The gradient does not exist at every point, leading to difficulties in defining the descent direction.*

• *Gradient usually does not exist at the optimal point.*

• *Difference approximation is not useful and may lead to serious failures.*

• *The (smooth) algorithm does not converge or it converges to a non-optimal point.*

Existence of critical points

We know that for a differentiable function f at $\bar{x}$ we have:

$$\bar{x} \text{ is a local minimum of } f \implies \nabla f(\bar{x}) = 0.$$

What happens if f is not differentiable?

Which tools can be used to replace differentiability?

Example 1: The distance function

Let C be a closed subset of some Banach space $(E, \|\cdot\|)$. The distance function of the set C is the function

$$x \mapsto d_C(x) := \inf_{u \in C} \|u - x\|.$$

$E = \mathbb{R}^n$

1. Suppose $\nabla d_C(x)$ exists and is different from 0. Then
   - x belongs to the complement of C.
   - There exists a unique point c in C closest to x.
   - $\nabla d_C(x) = \frac{x-c}{\|x-c\|}$.
2. Conversely, let $x \notin C$. If x has a unique closest point c in C, then d_C is differentiable at x and $\nabla d_C(x) = \frac{x-c}{\|x-c\|}$.

Problem of differentiabilty for $x \in C$?

Example 2: The minmax problem

The second problem concerned with nonsmoothness is the *minmax problem* :

$$\min g(x), \quad \text{where} \quad g(x) = \max_{u \in C} f(x,u) \tag{1}$$

where f is a smooth function with respect to x and C is a set.
The function g will not generally smooth even if f is.
A simple setting of this problem is the case where g is the maximum of two functions f_1 and f_2:

$$g(x) = \max(f_1(x), f_2(x)).$$

So the problem of nonsmoothness comes from the corner point $\bar{x}$ where $f_1(\bar{x}) = f_2(\bar{x})$.
Problem of differentiabilty to get "critical point condition"!!

The value function

Let $f : \mathbb{R}^n \times \mathbb{R}^p \mapsto \mathbb{R}$ be a function and $C \subset \mathbb{R}^p$ be a closed set. Define the function $v : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}$ by

$$v(x) = \inf_{y \in C} f(x, y).$$

Differentiability

If f is of class $\mathcal{C}^1$ and C is compact, then v is differentiable at $\bar{x}$ and $\nabla v(\bar{x}) = \nabla_x f(\bar{x}, \bar{y})$, where $\bar{y} \in C$ is a point satisfying $v(\bar{x}) = f(\bar{x}, \bar{y})$, provided that C and f are convex.

Example 3: Constrained optimization

Consider the following family of optimization problems

$$(P_\alpha) \begin{cases} \min f(x) \\ h(x) = \alpha \end{cases}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $h : \mathbb{R}^n \mapsto \mathbb{R}^p$ are given (smooth) functions. Let $v(\alpha)$ be the value of this problem. In general, this value function will take values in $[-\infty, +\infty]$. We have

$$f(x) \geq v(h(x)) \quad \forall x \in \mathbb{R}^n.$$

Necessary optimality

If $f(\bar{x}) = v(0)$, with $h(\bar{x}) = 0$, and v is differentiable at 0, then

$$\nabla f(\bar{x}) - Dh(\bar{x})\nabla v(0) = 0.$$

-$\nabla v(0)$ is a *Lagrange multiplier* at $\bar{x}$ for (P_0).

Problem of differentiabilty of v!!

Example 4: Constrained optimization: Penalization by the distance function

Consider the constrained optimization problem

$$\min f(x)$$
$$x \in A$$

Clarke penalization

Let A and B be closed sets in X, with $A \subset B$, and let $\bar{x} \in A$. Suppose that f is Lipschitz on B with constant K. Then the following assertions are equivalent :

1. $\bar{x}$ is a minimum of f over A,
2. For all $K' > K$, $\bar{x}$ is a minimum of the function

$$x \mapsto f(x) + K' d_A(x)$$

over B.

Problem of differentiabilty to get "critical point condition" !!

Example 5: Constrained optimization under calmness condition

Consider the constrained optimization problem

$$(P)\ \begin{array}{l}\min f(x)\\ g(x)\le 0\end{array}$$

f and g are continuous convex functions. Set $A := \{x \in E :\ g(x) \le 0\}$.

Penalaization under Calmness

Suppose that $\alpha := \inf\limits_{x\notin A} \dfrac{g(x)}{d_A(x)} > 0$ (Calmness condition). Suppose that f is K-Lipschitz. Then the following assertions are equivalent :

1. $\bar{x}$ is a solution of the problem (P),
2. $\forall \varepsilon \in]0, \alpha[$, $\bar{x}$ is a minimum of the function

$$x \mapsto f(x) + \frac{K}{\alpha - \varepsilon} g^+(x).$$

Constrained optimization leads to nondifferentiability

Example 6: Flow-Invariant sets

Let S be a closed subset of $\mathbb{R}^n$ and $\varphi : \mathbb{R}^n \mapsto \mathbb{R}^n$ be a locally Lipschitzian function. The question is whether the trajectories $x(t)$ of the differential equation

$$\dot{x}(t) = \varphi(x(t)), \quad x(0) = x_0 \tag{2}$$

leaves S invariant. In this case we say that the system (S, φ) is *flow-invariant*.

Example 6: Flow-Invariant sets

Let S be a closed subset of $\mathbb{R}^n$ and $\varphi : \mathbb{R}^n \mapsto \mathbb{R}^n$ be a locally Lipschitzian function. The question is whether the trajectories $x(t)$ of the differential equation

$$\dot{x}(t) = \varphi(x(t)), \quad x(0) = x_0 \tag{2}$$

leaves S invariant. In this case we say that the system (S, φ) is *flow-invariant*.

The set S is a smooth manifold if locally it admits a representation of the form

$$S = \{x \in \mathbb{R}^n : \ h(x) = 0\}$$

where $h : \mathbb{R}^n \mapsto \mathbb{R}^p$ is a continuously differentiable function with "nonvanishing" derivative on S.

Characterization

Let S be a smooth manifold. The system (2) is flow-invariant iff for every $x_0 \in S$, $\varphi(x_0)$ belong to the tangent space to S at x_0.

What happens if S is not smooth?

Example 7: Minimal time problem

By a *trajectory* of the standard control system

$$\dot{x} = f(x(t), u(t)) \text{ a.e.}, \quad u(t) \in U \text{ a.e.} \tag{3}$$

we mean a state function x corresponding to some choice of admissible (measurable) control function u. Here $f : \mathbb{R}^n \times \mathbb{R}^p \mapsto \mathbb{R}^n$ is a locally Lipschitzian mapping and $U \subset \mathbb{R}^p$ is a nonempty set. The *minimal time problem* refers to finding a trajectory that reach the origin as quickly as possible from a given point. The *minimal time function* T is defined on $\mathbb{R}^n$ by

$$T(\omega) = \inf\{T \geq 0 : \text{ some trajectory } x \text{ satisfies } x(0) = \omega,\ x(T) = 0\}.$$

The principe of optimality leads to: for any trajectory x,

$$s < t \quad \Longrightarrow \quad T(x(t)) - T(x(s)) \geq s - t,$$

that is, the function $\beta : t \mapsto T(x(t)) + t$ is increasing, and when x is optimal the function β is constant. So that we expect to have

$$\langle \nabla T(x(t)), \dot{x}(t) \rangle + 1 \geq 0$$

with equality when x is an optimal trajectory. The possible values of x for a trajectory being precisely the elements of the set $f(x(t), U)$, we arrive at

$$\min_{u \in U} \langle \nabla T(x(t)), f(x(t), u) \rangle + 1 = 0 \tag{4}$$

We define the (lower) Hamiltonian function h as follows:

$$h(x, p) := \min_{u \in U} \langle p, f(x(t), u) \rangle$$

In terms of h, the partial differential equation (4) above reads

$$h(x, \nabla T(x)) + 1 = 0 \tag{5}$$

a special case of the *Hamilton–Jacobi* equation.

The following questions arise :

Controllability: Is it always possible to steer ω to 0 in finite time?

Existence: do minimal-time trajectories exist?

Differentiability: **How do we know that T is differentiable?**

If this fails to be the case, then we shall need to replace the gradient ∇T used above by some suitably generalized derivative.

Example 8: Hamilton-Jacobi equations

These equations (of the first order type) are defined by mean of an Hamiltonian $H : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}$ as follows

$$\begin{cases} H(x, u(x), \nabla u(x)) = 0 & x \in \Omega \\ u = \varphi & \text{on } \partial\Omega \end{cases} \tag{6}$$

where $\Omega \subset \mathbb{R}^n$ is an open set and $\nabla u(x)$ denotes the gradient of u.

Example

Consider the following equation in $\mathbb{R}$

$$\begin{cases} |u'(x)| = 1 & x \in [0,1] \\ u(0) = 0, \quad u(1) = b & \end{cases}$$

- If $b = 0$: Many locally Lipschitzian solutions exist in the almost every where sense.
- If $0 \leq b < 1$: Existence of nonregular solutions
- If $b = 1$: Existence of a regular solution
- If $b > 1$: There is no continuous solution

Definition

A continuous function u is said to be a viscosity solution of the Hamilton-Jacobi equation (6) if

- $u = \varphi$ on $\partial\Omega$
- (Viscosity subsolution) for any $\mathcal{C}^\infty$-function v, if $u - v$ has a local maximum at $x_0 \in \Omega$, then

$$H(x_0, u(x_0), \nabla v(x_0)) \leq 0.$$

- (Viscosity supersolution) for any $\mathcal{C}^\infty$-function v, if $u - v$ has a local minimum at $x_0 \in \Omega$, then

$$H(x_0, u(x_0), \nabla v(x_0)) \geq 0.$$

Example 9:

Let $\Omega \subset \mathbb{R}^n$ be an open bounded set. Then the function $u = d(\cdot, \partial\Omega)$ is a viscosity solution of the following Hamilton-Jacobi equation

$$\begin{cases} |\nabla u| = 1 & \text{on } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

How to characterize these test functions ?

Example 10: Sweeping process

(An RCD (Residual Current Device) circuit).

Let us consider a circuit composed of a resistor R, a voltage source $u(t)$, an ideal diode, and a capacitor C mounted in series. The current through the circuit is denoted as $x(\cdot)$, and the charge of the capacitor is denoted as

$$z(t) = \int_0^t x(s)ds.$$

The dynamical equations are:

$$\begin{cases} \dot{z}(t) = -\frac{u(t)}{R} - \frac{1}{RC}z(t) + \frac{1}{R}v(t) \\ 0 \leq v(t) \perp w(t) := \frac{u(t)}{R} - \frac{1}{RC}z(t) + \frac{1}{R}v(t) \geq 0, \quad t \geq 0 \\ z(0) \in \mathbb{R}. \end{cases}$$

The last equation can be expressed in two manners :

- $v(t) = \max(0, -u(t) + \frac{1}{C}z(t))$.
- $v(t) = \text{proj}_{\mathbb{R}_+}(-u(t) + \frac{1}{C}z(t))$. (To be continued ...)

Some differentiability aspects

Let E be a Banach space. A function $f : E \mapsto \mathbb{R}$ is said to be

- Gâteaux differentiable at u if its directional derivative

$$Df(u) : h \mapsto f'(u,h) := \lim_{t \to 0^+} \frac{f(u+th) - f(u)}{t}$$

is linear and continuous

Some differentiability aspects

Let E be a Banach space. A function $f : E \mapsto \mathbb{R}$ is said to be

- Gâteaux differentiable at u if its directional derivative

$$Df(u) : h \mapsto f'(u, h) := \lim_{t \to 0^+} \frac{f(u + th) - f(u)}{t}$$

  is linear and continuous

- Fréchet differentiable at u if

$$\lim_{x \to u} \frac{f(x) - f(u) - Df(u)(x - u)}{\|x - u\|} = 0$$

Some differentiability aspects

Let E be a Banach space. A function $f : E \mapsto \mathbb{R}$ is said to be

- Gâteaux differentiable at u if its directional derivative

$$Df(u) : h \mapsto f'(u, h) := \lim_{t \to 0^+} \frac{f(u + th) - f(u)}{t}$$

  is linear and continuous

- Fréchet differentiable at u if

$$\lim_{x \to u} \frac{f(x) - f(u) - Df(u)(x - u)}{\|x - u\|} = 0$$

- Strictly differentiable at u if

$$\lim_{x, x' \to u} \frac{f(x) - f(x') - Df(u)(x - x')}{\|x - x'\|} = 0.$$

Elements of Nonsmooth Analysis

Making the parallel between smooth and nonsmooth objects

SMOOTH :	*NONSMOOTH :* $\forall h \in E$
1) *Gâteaux differentiability*	• $\lim\limits_{t \to 0^+} \dfrac{f(u+th) - f(x)}{t} \geq \langle x^*, h \rangle$
2) *Gâteaux differentiability*	• $\liminf\limits_{\substack{t \to 0^+ \\ h' \to h}} \dfrac{f(u+th') - f(u)}{t} \geq \langle x^*, h \rangle$
3) *Fréchet differentiability*	• $\liminf\limits_{x \to u} \dfrac{f(x) - f(u) - \langle x^*, x-u \rangle}{\|x-u\|} \geq 0$
4) *Strictly differentiability*	• $\limsup\limits_{\substack{x \to u \\ t \to 0^+}} \dfrac{f(x+th) - f(x)}{t} \geq \langle x^*, h \rangle$

Note that the introduction of non-smooth analysis tools depends largely on the geometry of the space considered.

Corresponding subdifferentials : Analytic construction

Fenchel subdifferential: Put $f'(x,h) := \lim_{t \to 0^+} \dfrac{f(x+tu) - f(x)}{t}$

$$\partial_{Fen} f(x) = \{x^* \in E^* : \langle x^*, h\rangle \le f'(x,h), \quad \forall h\}$$

Fréchet subdifferential:

$$\partial_F f(x) = \{x^* \in E^* : \liminf_{h \to 0} \frac{f(x+h) - f(x) - \langle x^*, h\rangle}{\| h \|} \ge 0\}$$

Dini subdifferential: Put $f^-(x,h) := \liminf_{\substack{u \to h \\ t \to 0^+}} \dfrac{f(x+tu) - f(x)}{t}$

$$\partial^- f(x) = \{x^* \in E^* : \langle x^*, h\rangle \le f^-(x,h), \quad \forall h\}$$

Clarke's subdifferential: Put $f^0(x_0,h) := \limsup_{\substack{x \to x_0 \\ t \to 0^+}} \dfrac{f(x+th) - f(x)}{t}$

$$\partial_c f(x_0) = \{x^* \in E^* : \langle x^*, h\rangle \le f^0(x_0,h), \quad \forall h\}$$

Limiting contrepart

Limiting subdifferential:

$$\partial_L f(x_0) = w^* - seq - \limsup_{x \xrightarrow{f} x_0} \partial_F f(x).$$

Sequential approximate subdifferential:

$$\partial_A^{\text{seq}} f(x_0) = w^* - seq - \limsup_{x \xrightarrow{f} x_0} \partial^- f(x).$$

Relationships for locally Lipschitz functions

Asplund spaces :

$$\partial_C f(x_0) = \mathrm{cl}^* \ \mathrm{co}\partial_L f(x_0) = \mathrm{cl}^* \ \mathrm{co}\partial_A^{\mathrm{seq}} f(x_0)$$

WCG Asplund spaces :

$$\partial_L f(x_0) = \partial_A^{\mathrm{seq}} f(x_0)$$

Outside Asplund spaces

$$\partial_C f(x_0) \supseteq \mathrm{cl}^* \ \mathrm{co}\partial_A^{\mathrm{seq}} f(x_0) \supsetneq \mathrm{cl}^* \ \mathrm{co}\partial_L f(x_0)$$

Asplund spaces: Banach spaces on which every continuous convex function is Fréchet differentiable at a dense set of points.

Chain rules

$f, g : E \mapsto \mathbb{R} \cup \{+\infty\}$ lsc, with g locally Lipschitzian at x_0.

$$\partial_{Fen}(f+g)(x_0) = \partial_{Fen} f(x_0) + \partial_{Fen} g(x_0) \quad (f, g \text{ convex})$$

$$\partial_C(f+g)(x_0) \subseteq \partial_C f(x_0) + \partial_C g(x_0)$$

$$\partial_A(f+g)(x_0) \subseteq \partial_A f(x_0) + \partial_A g(x_0)$$

$$\partial_L(f+g)(x_0) \not\subseteq \partial_L f(x_0) + \partial_L g(x_0)$$

where $\partial_A f(x)$ denotes the (topological)approximate subdifferential

Asplund spaces:

$$\partial_L(f+g)(x_0) \subset \partial_L f(x_0) + \partial_L g(x_0)$$

Example

Let $X = L^1[0,1]$ and let $f(u) = \int_0^1 |\sin u(t)| dt$ and $g = -f$. Then $\partial_L(f+g)(0) = \{0\}$ while $\partial_L f(0) = \{0\}$ and $\partial_L g(0) = \emptyset$.

Fenchel subdifferential

The Fenchel Subdifferential

Definition.

The subdifferential of a convex function $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ at x in the effective domain $\mathrm{Dom} f$ of f is the set

$$\partial_{Fen} f(x) = \Big\{x^* \in E^* : \langle x^*, y - x\rangle + f(x) \leq f(y) \quad \forall y \in E\Big\}.$$

Each vector $x^* \in \partial_{Fen} f(x)$ is called a subgradient of f at x.

Characterization of critical points.

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a convex function. Then the following are equivalent:

- $\bar{x}$ is a local minimum of f.
- $\bar{x}$ is a global minimum of f.
- $0 \in \partial_{Fen} f(\bar{x})$.

Geometric representation.

Normal cone.

Let $C \subset E$ be a closed convex set containing $\bar{x}$. The normal cone to C at $\bar{x}$ is the weak-star closed convex cone defined by

$$N_{Fen}(C,\bar{x}) = \{x^* \in E^* : \langle x^*, x - \bar{x}\rangle \leq 0 \, \forall x \in C\},$$

that is, $N_{Fen}(C,\bar{x}) = \partial \Psi_C(\bar{x})$, where Ψ_C is the indicator function of C.

Geometric representation.

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a lsc convex function and let $\bar{x} \in \mathrm{Dom} f$. Then

$$\partial_{Fen} f(\bar{x}) = \left\{x^* \in E^* : (x^*, -1) \in N_{Fen}(\mathrm{epi} f, (\bar{x}, f(\bar{x}))\right\},$$

where $\mathrm{epi} f$ is the epigraph of f.

Weak-star compactness.

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a convex function, which is continuous and finite at $\bar{x}$. Then $\partial_{Fen} f(\bar{x})$ is a non empty convex and w^*-compact set.

Directional derivative.

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a convex function. Then the classical directional derivative $f'(x; d)$ exists in every direction $d \in E$ and for all $x \in \text{Dom} f$

- $\partial_{Fen} f(x) = \left\{x^* \in E^* : \langle x^*, d\rangle \leq f'(x; d) \, \forall d \in E\right\}$
- when f is continuous at x,

$$f'(x; d) = \max_{x^* \in \partial_{Fen} f(x)} \langle x^*, d\rangle \, \forall d \in E.$$

Chain rules.

Let $f, g : E \mapsto \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous convex functions finite at $\bar{x} \in E$. If g is continuous at $\bar{x}$, then

$$\partial_{Fen}(f + g)(\bar{x}) = \partial_{Fen} f(\bar{x}) + \partial_{Fen} g(\bar{x}).$$

Clarke subdifferential

Clarke subdifferential : The analytic contruction.

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a **locally Lipschiltzian** function at $\bar{x}$.

Clarke directional derivative.

The Clarke directional derivative of f at $\bar{x}$ in the direction $d \in E$ is defined by

$$f^\circ(\bar{x}, d) = \limsup_{\substack{x \to \bar{x} \\ t \to 0^+}} \frac{f(x+td) - f(x)}{t}.$$

Some properties.

- The function $d \mapsto f^\circ(\bar{x}, d)$ is positively homogeneous, convex and continuous.
- For all $d \in E$, $f^\circ(\bar{x}, -d) = (-f)^\circ(\bar{x}, d)$.

The analytic construction.

The Clarke subdifferential of f at $\bar{x}$ is the w^*-compact convex set defined by

$$\partial_c f(\bar{x}) = \{x^* \in E^* : \langle x^*, d \rangle \leq f^\circ(\bar{x}, d) \forall d \in E\}.$$

Clarke tangent and normal cones

Let $C \subset E$ be a closed set containing $\bar{x}$.

Clarke tangent cone.

The Clarke tangent cone of C at $\bar{x}$ is the closed convex cone given by

$$T_c(C,\bar{x}) := \liminf_{\substack{x \xrightarrow{C} \bar{x} \\ t \to 0^+}} \frac{C - x}{t} = \left\{ h \in E : \; d_C^{\circ}(\bar{x}, h) = 0 \right\}.$$

Clarke normal cone.

The Clarke normal cone to C at $\bar{x}$ is the w^*-closed convex cone given by
$$N_c(C,\bar{x}) = \{x^* \in E^* : \langle x^*, h\rangle \leq 0 \, \forall h \in T_c(C,\bar{x})\}.$$

Clarke normal cone.

$$N_c(C,\bar{x}) = \mathrm{cl}^* \mathrm{cone}\big(\partial_c d_C(\bar{x})\big).$$

Let $C \subset \mathbb{R}^n$ be a closed set containing $\bar{x}$.

Let $C \subset \mathbb{R}^n$ be a closed set containing $\bar{x}$.

Penot (1981), Cornet (1981)

$$T_c(C,\bar{x}) = \liminf_{x \xrightarrow{C} \bar{x}} K(C,x)$$

where $K(C,x)$ denotes the contingent cone to C at x.

Let $C \subset \mathbb{R}^n$ be a closed set containing $\bar{x}$.

Penot (1981), Cornet (1981)

$$T_c(C, \bar{x}) = \liminf_{x \xrightarrow{C} \bar{x}} K(C, x)$$

where $K(C, x)$ denotes the contingent cone to C at x.

$$d_{T_c(C,\bar{x})}(h) \geq d_C^0(\bar{x}, h) \quad \forall h$$

or equivalently

$$\partial_c d_C(\bar{x}) \subset \mathbb{B} \cap N_c(C, \bar{x}).$$

Let $C \subset \mathbb{R}^n$ be a closed set containing $\bar{x}$.

Penot (1981), Cornet (1981)

$$T_c(C, \bar{x}) = \liminf_{x \overset{C}{\to} \bar{x}} K(C, x)$$

where $K(C, x)$ denotes the contingent cone to C at x.

$$d_{T_c(C,\bar{x})}(h) \geq d_C^0(\bar{x}, h) \quad \forall h$$

or equivalently

$$\partial_c d_C(\bar{x}) \subset \mathbb{B} \cap N_c(C, \bar{x}).$$

Watkins (1985)

For any $v \in T_c(C, \bar{x})$ and any real number $\ell > \|v\|$, there exists a Lipschitz continuous mapping $c : [0, 1] \mapsto C$ with Lipschitz constant ℓ such that c is strictly right differentiable at 0 with $c(0) = \bar{x}$ and $c'(0) = v$.

Geometric construction of the Clarke subdifferential.

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a lsc function and let $\bar{x} \in \text{Dom} f$. The Clarke subdifferential of f at $\bar{x}$ is the w^*-closed and convex set defined by

$$\partial_c f(\bar{x}) = \left\{x^* \in E^* : (x^*, -1) \in N_c(\text{epi} f, (\bar{x}, f(\bar{x}))\right\}.$$

Clarke directional derivative as a support function.

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a locally Lipschitzian function at $\bar{x}$. Then

$$f^\circ(\bar{x}; d) = \max_{x^* \in \partial_c f(\bar{x})} \langle x^*, d \rangle \, \forall d \in E.$$

Rademacher Theorem.

Let $S \subset \mathbb{R}^n$ be an open set. A function $f : S \mapsto \mathbb{R}$ that is locally Lipschitz on S is differentiable almost everywhere on S.

This leads to the following construction of the Clarke subdifferential.

Gradient characterization of the Clarke subdifferential.

Let $f : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}$ be a locally Lipschitzian function at $\bar{x}$. Then

$$\partial_c f(\bar{x}) = \text{cl}^*\text{conv}\Big\{x^* \in \mathbb{R}^n : \nabla f(x_i) \to x^*, x_i \to \bar{x} \text{ and } f \text{ is differentiable at } x_i\Big\},$$

where $\text{cl}^*\text{conv}S$ denotes the w^*-closed convex hull of the set S.

Sum

Let $f_1, f_2 : E \mapsto \mathbb{R} \cup \{+\infty\}$ be lsc functions and $\bar{x} \in \text{Dom} f_1 \cap \text{Dom} f_2$ with f_1 locally Lipschitz at $\bar{x}$. Then

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

Subdifferential of the maximum function

Let $f_1, \cdots, f_n : E \mapsto \mathbb{R}$ be locally Lipschitzian function around $\bar{x}$, with $f_1(\bar{x}) = \cdots = f_n(\bar{x})$. Then

$$\partial_c(\max_{i=1,\cdots,n} f_i)(\bar{x}) \subset \text{co}[\partial_c f_i(\bar{x}), i = 1, \cdots, n].$$

The equality holds whenever all f_i are Clarke regular at $\bar{x}$.

Clarke subdifferential of the distance function.

Let c belong to $C \subset \mathbb{R}^n$. Then

$$\partial d_C(c) = \text{co}\{0, \lim_{i\to+\infty} \frac{x_i - c_i}{\|x_i - c_i\|}\},$$

where we consider all sequences (x_i), (c_i) such that x_i is not in C and has closest point (c_i) in C, and $\lim_{i\to+\infty} x_i = c$.

Exercise.

Compute the subdifferential of the distance function of the set $C = \{(x, y) \in \mathbb{R}^2 : xy = 0\}$ at $(0, 0)$.

When Clarke's subdifferential is a singleton?

f is Hadamard strictly differentiable (HSD) at u if there exists $D_s f(u) \in E^*$ such that

$$\lim_{\substack{x \to u \\ t \to 0^+}} \frac{f(x+th) - f(x)}{t} = \langle D_s f(u), h \rangle$$

and provided the convergence is uniform for h in compact sets.

Characterization of the HSD

Let $x^* \in E^*$. Then the following assertions are equivalent:

1. f is HSD at u with $D_s f(u) = x^*$.
2. f is Lipschitz near u and

$$\lim_{\substack{x \to u \\ t \to 0^+}} \frac{f(x+th) - f(x)}{t} = \langle x^*, h \rangle$$

When Clarke's subdifferential is a singleton?

f is Hadamard strictly differentiable (HSD) at u if there exists $D_s f(u) \in E^*$ such that

$$\lim_{\substack{x \to u \\ t \to 0^+}} \frac{f(x+th) - f(x)}{t} = \langle D_s f(u), h \rangle$$

and provided the convergence is uniform for h in compact sets.

Characterization of the HSD

Let $x^* \in E^*$. Then the following assertions are equivalent:

1. f is HSD at u with $D_s f(u) = x^*$.
2. f is Lipschitz near u and

$$\lim_{\substack{x \to u \\ t \to 0^+}} \frac{f(x+th) - f(x)}{t} = \langle x^*, h \rangle$$

Uniqueness of the Clarke's subgradient

Let f Lipschitz near u. Then f is HSD at u IFF $\partial_C f(u)$ is a singleton.

Proximal and limiting proximal subdifferentials

Metric projection.

Let E be a **Hilbert space** and $C \subset E$ be a closed nonempty set. The metric projection onto C is the set-valued mapping $P_C : E \rightrightarrows E$ defined by

$$P_C(x) = \{u \in C : \, d_C(x) = \|x - u\|\}.$$

Some properties of the metric projection.

Let $x \in E$ and $u \in C$. Then the following assertions are equivalent:

- $u \in P_C(x)$;
- $u \in P_C(u + t(x - u)) \ \forall t \in [0, 1]$;
- $d_C(u + t(x - u)) = t\|x - u\| \ \forall t \in [0, 1]$;
- $\langle x - u, u' - u \rangle \leq \frac{1}{2}\|u' - u\|^2 \ \forall u' \in C$.

Proximal normal cones.

Let $\bar{x} \in C$

Proximal normal cone.

The proximal normal cone to C at $\bar{x}$ is the convex cone given by

$$N_p(C, \bar{x}) := \mathrm{cone}\big(P_C^{-1}(\bar{x}) - \bar{x}\big).$$

Variational characterization of the proximal normal cone.

$$N_p(C, \bar{x}) = \{x^* \in E : \exists \alpha > 0; \langle x^*, x - \bar{x} \rangle \leq \alpha \|x - \bar{x}\|^2 \, \forall x \in C\}.$$

In fact the notion of proximal normals is essentially a local property.

Local characterization.

For any given $\delta > 0$, we have $x^* \in N_p(C, \bar{x})$ iff there exists $\alpha > 0$ such that

$$\langle x^*, x - \bar{x} \rangle \leq \alpha \|x - \bar{x}\|^2 \, \forall x \in C \cap B(\bar{x}, \delta).$$

Proximal normal to smooth sets.

Suppose that C has the following representation:
$C = \{x \in E : h_i(x) = 0, i = 1, \cdots, k\}$ where $h_i : E \mapsto \mathbb{R}$ is $\mathcal{C}^1$. If $\{\nabla h_i(\bar{x}), i = 1, \cdots, k\}$ are linearly independent, then

- $N_p(C, \bar{x}) \subset \operatorname{span}\{\nabla h_1(\bar{x}), \cdots, \nabla h_k(\bar{x})\}$.
- If in addition each h_i is $\mathcal{C}^2$, then the equality holds.

Proximal subdifferential of the distance function.

Let $x \notin C$ and $x^* \in \partial_p d_C(x)$. Then there exists $u \in C$ so that:

- Every minimizing sequence $(u_i) \subset C$ of $\inf_{v \in C} \|x - v\|$ converges to u.
- $P_C(x) = \{u\}$.
- The Fréchet derivative $\nabla d_C(x)$ exists, and

$$\{x^*\} = \partial_p d_C(x) = \{\nabla d_C(x)\} = \Big\{\frac{x - u}{\|x - u\|}\Big\}.$$

- $x^* \in N_p(C, u)$.

Proximal subdifferential: Geometric and variational characterizations.

Geometric characterization.

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a lsc function and let $\bar{x} \in \mathrm{Dom} f$. The proximal subdifferential is defined by

$$\partial_p f(\bar{x}) = \left\{ x^* \in E : \, (x^*, -1) \in N_p(\mathrm{epi} f, (\bar{x}, f(\bar{x}))\right\}.$$

Variational characterization.

The following assertions are equivalent:

- $x^* \in \partial_p f(\bar{x})$.
- $\exists \alpha > 0$ and $\delta > 0$ such that

$$f(x) - f(\bar{x}) \geq \langle x^*, x - \bar{x} \rangle - \alpha \|x - \bar{x}\|^2 \, \forall x \in B(\bar{x}, \delta).$$

Link with Gâteaux derivative.

Assume that f is Gâteaux differentiable at $\bar{x}$. Then

- $\partial_p f(\bar{x}) \subset \{\nabla f(\bar{x})\}$.
- If f is of class $\mathcal{C}^2$ is some neighbourhood V of $\bar{x}$, then $\partial_p f(x) \subset \{\nabla f(x)\}\, \forall x \in V$.

The first inclusion may be strict even if f is of class $\mathcal{C}^1$. To see this, take $f(x) = -\sqrt{|x|^3}$ and $\bar{x} = 0$. But we have the following:

Density theorem

Let $\bar{x} \in \mathrm{Dom} f$, where f is lsc, and let $\varepsilon > 0$ be given. Then there exists $x \in B(\bar{x}, \varepsilon)$, with $f(\bar{x}) - \varepsilon \leq f(x) \leq f(\bar{x})$, such that $\partial_p f(x) \neq \emptyset$. In particular $\mathrm{Dom}\partial_p f$ is dense in $\mathrm{Dom} f$.

Optimality conditions.

- If f has a local minimum at $\bar{x}$, then $0 \in \partial_p f(\bar{x})$.
- Let f of class $\mathcal{C}^2$, and suppose that $\bar{x}$ is a local minimum of f over C. Then $-\nabla f(\bar{x}) \in N_p(C, \bar{x})$.

Limiting counterpart.

Limiting proximal normal.

The limiting proximal normal cone to C at $\bar{x}$ is the set

$$N_\ell(C,\bar{x}) = w - seq - \limsup_{x \xrightarrow{C} \bar{x}} N_p(C,x).$$

Before giving a geometric and variational characterizations of the limiting proximal subdifferentials, let us point out that, unlike the proximal normal cone, the limiting one is not convex. To see this consider
$C := \{(x_1,x_2) \in \mathbb{R}^2 : \; x_2 \geq -|x_1|\}$. Then
$N_\ell(C,(0,0)) = \{(y,y) : \; y \leq 0\} \cup \{(y,-y) : \; y \geq 0\}$.

Limiting proximal subdifferential: Geometric characterization

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a lsc function and $\bar{x} \in \mathrm{Dom} f$. The limiting proximal subdifferential of f at $\bar{x}$ is given by

$$\partial_\ell f(\bar{x}) = \left\{x^* \in E^* : \; (x^*,-1) \in N_\ell(\mathrm{epi} f, (\bar{x}, f(\bar{x})))\right\}.$$

Proximal subdifferential: Analytic construction. Chain rule.

Proximal subdifferential: Analytic construction.

$$\partial_\ell f(\bar{x}) = w - seq - \limsup_{x \xrightarrow{f} \bar{x}} \partial_p f(x).$$

Sum.

Let $f_1, f_2 : E \mapsto \mathbb{R} \cup \{+\infty\}$ be lsc functions and $\bar{x} \in \mathrm{Dom} f_1 \cap \mathrm{Dom} f_2$ with f_1 locally Lipschitz at $\bar{x}$. Then

- For any $x^* \in \partial_p(f_1 + f_2)(\bar{x})$ and $\varepsilon > 0$ there exist (for $i = 1, 2$) points $x_i \in B(\bar{x}, \varepsilon)$, with $|f(x_i) - f(\bar{x})| < \varepsilon$, such that

$$x^* \in \partial_p f_1(x_1) + \partial_p f_2(x_2) + B(0, \varepsilon).$$

- $\partial_\ell(f_1 + f_2)(\bar{x}) \subset \partial_\ell f_1(\bar{x}) + \partial_\ell f_2(\bar{x})$.

Composition.

Let F be a Hilbert space and $F : E \mapsto F$ and $g : F \mapsto \mathbb{R}$ be locally Lipschitzian mappings at $\bar{x}$ and $F(\bar{x})$ respectively. Then

- For any $x^* \in \partial_p(g \circ F)(\bar{x})$ and $\varepsilon > 0$ there exist $x \in B(\bar{x}, \varepsilon)$, with $\|F(x) - F(\bar{x})\| < \varepsilon$, $y \in B(F(\bar{x}, \varepsilon)$ and $y^* \in \partial_p g(y)$ such that

$$x^* \in \partial_p\big(\langle y^*, F(\cdot)\rangle\big)(x) + B(0, \varepsilon).$$

- If $\dim Y < \infty$, then

$$\partial_\ell(g \circ F)(\bar{x}) \subset \bigcup_{y^* \in \partial_\ell g(F(\bar{x}))} \partial_\ell\big(\langle y^*, F(\cdot)\rangle\big)(\bar{x}).$$

Normal cones via subdifferential of the distance function

Proximal normal cone.

Assume that E is a Hilbert space. Then

$$N_p(C, \bar{x}) = \text{cone}\big(\partial_p d_C(\bar{x})\big).$$

Limiting proximal normal cone.

Assume that E is a Hilbert space. Then

$$N_\ell(C, \bar{x}) = \text{clcone}\big(\partial_\ell d_C(\bar{x})\big).$$

Relationship between Clarke subdifferential and limiting proximal subdifferential

For locally Lipschitzian function.

Assume that E is a Hilbert space and f is locally Lipschitz at $\bar{x}$. Then

$$\partial_C f(\bar{x}) = \mathrm{cl}^* \, \mathrm{co} \partial_\ell f(\bar{x}).$$

Fréchet and limiting Fréchet subdifferentials

Limiting Fréchet subdifferential.

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a lsc function and let $\bar{x} \in \mathrm{Dom} f$.

Fréchet subdifferential.

Let $\varepsilon \geq 0$. The $\varepsilon-$Fréchet subdifferential of F at $\bar{x}$ is the convex and norm-closed set

$$\partial_F^{\varepsilon} f(\bar{x}) = \{x^* : \liminf_{h \to 0} \frac{f(\bar{x}+h) - f(\bar{x}) - \langle x^*, h\rangle}{\| h \|} \geq -\varepsilon\}.$$

For $\varepsilon = 0$, we put $\partial_F f(\bar{x}) = \partial_F^0 f(\bar{x})$.

Limiting Fréchet subdifferential.

The limiting Fréchet subdifferential of f at $\bar{x}$ is the set

$$\partial_L f(\bar{x}) = w^* - seq - \limsup_{\substack{x \xrightarrow{f} \bar{x} \\ \varepsilon \to 0^+}} \partial_F^{\varepsilon} f(x).$$

Asplund spaces *are Banach spaces on which every continuous convex function is Fréchet differentiable at a dense set of points.*

Characterizations of Asplund spaces.

Characterizations.

The following assertions are equivalent:

1. E is Asplund.
2. For any $\varepsilon \geq 0, \delta > 0, \gamma > 0$ and any lsc functions $f_1, f_2 : E \mapsto \mathbb{R} \cup \{+\infty\}$ and $\bar{x} \in \mathrm{Dom} f_1 \cap \mathrm{Dom} f_2$ with f_1 locally Lipschitz at $\bar{x}$
$$\begin{aligned}\partial_F^{\varepsilon}(f_1 + f_2)(\bar{x}) \subset \quad & \big\{\partial_F f_1(x_1) + \partial_F f(x_2) : x_i \in B(\bar{x}, \delta), \\ & |f_i(x_i) - f_i(\bar{x})| < \delta, i = 1, 2\big\} + (\varepsilon + \gamma) B_{E^*}.\end{aligned}$$
3. For any lsc function $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ and $\bar{x} \in \mathrm{Dom} f$
$$\partial_L f(\bar{x}) = w^* - seq - \limsup_{x \xrightarrow{f} \bar{x}} \partial_F f(x).$$
4. For any lsc functions $f_1, f_2 : E \mapsto \mathbb{R} \cup \{+\infty\}$ and $\bar{x} \in \mathrm{Dom} f_1 \cap \mathrm{Dom} f_2$ with f_1 locally Lipschitz at $\bar{x}$
$$\partial_L(f_1 + f_2)(\bar{x}) \subset \partial_L f_1(\bar{x}) + \partial_L f_2(\bar{x}).$$

Normal cones.

Let $C \subset E$ be a closed set containing $\bar{x}$.

Fréchet normal cone.

Let $\varepsilon \geq 0$. The $\varepsilon-$Fréchet normal cone to C at $\bar{x}$ is the set

$$N_F^{\varepsilon}(C,\bar{x}) = \partial_F^{\varepsilon}\Psi_C(\bar{x}) = \{x^* \in E^* : \limsup_{x \xrightarrow{C} \bar{x}} \frac{\langle x^*, x-\bar{x}\rangle}{\|x-\bar{x}\|} \leq \varepsilon\}.$$

Limiting Fréchet normal cone.

The limiting Fréchet normal cone to C at $\bar{x}$ is the set

$$N_L(C,\bar{x}) = w^* - seq - \limsup_{\substack{x \xrightarrow{C} \bar{x} \\ \varepsilon \to 0^+}} N_F^{\varepsilon}(C,x).$$

Normal cone characterization of Asplund spaces.

A characterization of Asplund spaces.

The following assertions are equivalent:

1. E is Asplund.
2. For any closed set $C \subset E$ and any boundary point $\bar{x} \in C$

$$N_L(C, \bar{x}) = w^* - seq - \limsup_{x \xrightarrow{C} \bar{x}} N_F(C, x).$$

Remarks.

Remark.

Note that in the finite dimensional case $E = \mathbb{R}^n$, the limiting Fréchet normal cone coincides with the one in Mordukhovich [13]:

$$N_p(C, \bar{x}) == \limsup_{x \to \bar{x}} \operatorname{cone}\big(x - P_C(x)\big)$$

where "cone" stands for the conic hull of a set and $P_C(x)$ means the Euclidean projection of x on the closure of C.

Remark

- The limiting Fréchet normal cone is not convex.
- There are a closed subset C of the Hilbert space ℓ^2 and a boundary point $\bar{x} \in C$ such that $N_L(C, \bar{x})$ is not norm closed.

Geometric characterization.

Geometric characterization.

Let $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a lsc function and $\bar{x} \in \mathrm{Dom} f$. The limiting Fréchet subdifferential of f at $\bar{x}$ is given by

$$\partial_L f(\bar{x}) = \big\{x^* \in E^* : (x^*, -1) \in N_L(\mathrm{epi} f, (\bar{x}, f(\bar{x}))\big\}.$$

Density theorem

The density theorem and the Ekeland variational principle

In this section, ∂ will denote one of the previous subdifferentials with appropriate Banach spaces.

(Density theorem).

Let $\bar{x} \in \mathrm{Dom} f$, and let $\varepsilon > 0$ be given. Then there exists $x \in B(\bar{x}, \varepsilon)$, with $f(\bar{x}) - \varepsilon \leq f(x) \leq f(\bar{x})$, such that $\partial f(x) \neq \emptyset$. In particular $\mathrm{Dom}\partial f$ is dense in $\mathrm{Dom} f$.

The density theorem and the Ekeland variational principle

In this section, ∂ will denote one of the previous subdifferentials with appropriate Banach spaces.

(Density theorem).

Let $\bar{x} \in \mathrm{Dom} f$, and let $\varepsilon > 0$ be given. Then there exists $x \in B(\bar{x}, \varepsilon)$, with $f(\bar{x}) - \varepsilon \leq f(x) \leq f(\bar{x})$, such that $\partial f(x) \neq \emptyset$. In particular $\mathrm{Dom}\partial f$ is dense in $\mathrm{Dom} f$.

Ekeland variational principle

Let E be a Banach space and $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ be a lsc function which is bounded from below on some closed set $C \subset E$. Then, given $\gamma > 0$, and $u \in C$, with $f(u) < +\infty$, there exists $v \in C$ such that

$$f(v) \leq f(x) + \gamma \|x - v\| \quad \forall x \in C$$

$$f(v) + \gamma \|u - v\| \leq f(u).$$

Applications

- Classical Analysis
- Optimization
- Optimal control

Classical Analysis

Nonsmooth classical analysis

Lebourg mean value theorem.

Suppose that f is locally Lipschitz on some open convex set Ω. For each $a, b \in \Omega$, with $a \neq b$ there exists $c \in [a; b)$ and $x^* \in \partial_c f(c)$ such that

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

Zagrodny mean value theorem

Given a lsc function $f : E \mapsto \mathbb{R} \cup \{+\infty\}$, for each $a, b \in \mathrm{Dom} f$, with $a \neq b$ there exists $c \in [a; b)$ and two sequences

- $(x_k) \subset E$, $\lim\limits_{k \to +\infty} x_k = c$;
- $(x_k^*) \subset E^*$, with $x_k^* \in \partial_c f(x_k)$

such that

$$\liminf_{k \to +\infty} \langle x_k^*, b - a \rangle \geq f(b) - f(a)$$

$$\liminf_{k \to +\infty} \langle x_k^*, b - x_k \rangle \geq \frac{\|b - c\|}{\|b - a\|} (f(b) - f(a)).$$

Given a lsc function $f : E \mapsto \mathbb{R} \cup \{+\infty\}$. The lower Dini derivative of f at $x \in \text{Dom} f$ in the direction $d \in E$ is

$$d^- f(x, d) = \liminf_{t \to 0^+} \frac{f(x + td) - f(x)}{t}.$$

Diewert mean value theorem.

For each $a, b \in \text{Dom} f$, with $a \neq b$ there exists $c \in [a; b)$ such that

$$d^- f(c, b - a) \geq f(b) - f(a).$$

Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a function. One says that f is decreasing if

$$f(y) \leq f(x) \text{ whenever } x \leq y.$$

The inequality $x \leq y$ is understood in the component-wise sense: $x_i \leq y_i$, $i = 1, \cdots, n$.

Characterization of the monotony.

f is decreasing iff $x^* \leq 0 \ \forall x^* \in \partial_c f(x)$, $\forall x \in \mathbb{R}^n$.

Extension

Let $K \subset E$ be a closed convex cone and K^0 be the negative polar of K, that is,

$$K^0 := \{x^* \in E^* : \langle x^*, h \rangle \leq 0 \, \forall h \in K\}.$$

Characterization of the monotony.

Let $f : E \mapsto \mathbb{R}$ be a function. The following assertions are equivalent:

- f is decreasing with respect to K, that is for all $x, y \in E$, with $y - x \in K$, $f(y) \leq f(x)$,
- $\partial_c f(x) \subset K^0$, $\forall x \in E$.

Optimization

Necessary optimality conditions.

- E and F are Hilbert spaces
- $f : E \mapsto \mathbb{R} \cup \{+\infty\}$ and $g : E \mapsto F$ are mappings
- $C \subset E$ and $D \subset F$ are closed sets

The optimization problem :

$$(P) \quad \begin{cases} \min f(x) \\ x \in C \\ g(x) \in D \end{cases}$$

In this part of applications, we use the notation ∂ for the limiting Proximal subdifferential or for the Clarke subdifferential.

Lagrange multipliers

Fritz-John Lagrange multipliers

(λ, z^*) is a Fritz-John Lagrange multiplier for (P) at $\bar{x}$ if

FJ1 $(\lambda, z^*) \neq 0$

FJ2 $\lambda \geq 0$, $z^* \in N(D, g(\bar{x}))$

L2 $0 \in \lambda \partial f(\bar{x}) + \partial(z^* \circ g)(\bar{x}) + N(C, \bar{x})$

Fritz-John multipliers with $Y = \mathbb{R}^m$ or D is a closed convex cone with nonempty interior

Suppose that f and g are locally Lipschitz at $\bar{x}$ local solution for (P). Then there exists a Fritz-John Lagrange multiplier (λ, z^*) for (P) at $\bar{x}$.

The failure of the necessary optimality conditions

Brokate

Let $X = Y = l^2$ be the Hilbert space of square summable sequences, with (e_k) its canonical orthonormal base and let the operator $A : l^2 \to l^2$ be defined by

$$A(\sum x_i e_i) = \sum 2^{1-i} x_i e_i.$$

Then A is not surjective and $\mathrm{Im}(A)$ is a proper dense subspace of l^2. The adjoint A^* is injectif but not surjectif. So let $x^* \notin \mathrm{Im}(A^*)$ and set $f = x^*$, $g = A$ and $D = \{0\}$. Then 0 is only the feasible point and it is the optimum for this problem. Moreover there is no Fritz-John Lagrange multiplier for this problem at 0.

The failure of the necessary optimality conditions

Brokate

Let $X = Y = l^2$ be the Hilbert space of square summable sequences, with (e_k) its canonical orthonormal base and let the operator $A : l^2 \to l^2$ be defined by

$$A(\sum x_i e_i) = \sum 2^{1-i} x_i e_i.$$

Then A is not surjective and $\mathrm{Im}(A)$ is a proper dense subspace of l^2. The adjoint A^* is injectif but not surjectif. So let $x^* \notin \mathrm{Im}(A^*)$ and set $f = x^*$, $g = A$ and $D = \{0\}$. Then 0 is only the feasible point and it is the optimum for this problem. Moreover there is no Fritz-John Lagrange multiplier for this problem at 0.

What is missing here?

The failure of the necessary optimality conditions

Brokate

Let $X = Y = l^2$ be the Hilbert space of square summable sequences, with (e_k) its canonical orthonormal base and let the operator $A : l^2 \to l^2$ be defined by

$$A(\sum x_i e_i) = \sum 2^{1-i} x_i e_i.$$

Then A is not surjective and $\text{Im}(A)$ is a proper dense subspace of l^2. The adjoint A^* is injectif but not surjectif. So let $x^* \notin \text{Im}(A^*)$ and set $f = x^*$, $g = A$ and $D = \{0\}$. Then 0 is only the feasible point and it is the optimum for this problem. Moreover there is no Fritz-John Lagrange multiplier for this problem at 0.

What is missing here? Closedness of the rang of A

When do we get Fritz-John-Lagrange multipliers if $\dim Y = +\infty$?

Before the 90', the only well known results when $\dim Y = +\infty$ assumed that

D is a closed convex cone with $\mathrm{int} D \neq \emptyset$.

Fritz-John multipliers

Let $\bar{x}$ be a solution of the problem (P) at f is locally Lipschitzian and g is of class $\mathcal{C}^1$. Suppose D is a closed convex cone with non empty interior. Then there exist $\lambda \geq 0$ and $y^* \in N(D, g(\bar{x}))$, with $(\lambda, y^*) \neq 0$, such that

$$0 \in \lambda \partial f(\bar{x}) + D^* g(\bar{x}(y^*) + N(C, \bar{x}).$$

How to avoid the interiority assumption?

How to include the finite-dimensional situation?

Fritz-John multipliers

Let $\bar{x}$ be a solution of the problem (P) at f is locally Lipschitzian and g is of class $\mathcal{C}^1$. Suppose D is a closed convex cone with non empty interior. Then there exist $\lambda \geq 0$ and $y^* \in N(D, g(\bar{x}))$, with $(\lambda, y^*) \neq 0$, such that

$$0 \in \lambda \partial f(\bar{x}) + D^* g(\bar{x}(y^*) + N(C, \bar{x}).$$

How to avoid the interiority assumption?

How to include the finite-dimensional situation?

The answers to these questions are given in J. and Thibault (1995) where the unification appears for the first time.

Existence of Karush-Kuhn-Tucker(KKT) multipliers

KKT multipliers

z^* is a KKT Lagrange multiplier for (P) at $\bar{x}$ if

L1 $z^* \in N(D, g(\bar{x}))$

L2 $0 \in \partial f(\bar{x}) + \partial(z^* \circ g)(\bar{x}) + N(C, \bar{x})$

Let $KKT(\bar{x})$ denotes the set of KKT Lagrange multiplier for (P) at $\bar{x}$.

Existence of Karush-Kuhn-Tucker(KKT) multipliers

KKT multipliers

z^* is a KKT Lagrange multiplier for (P) at $\bar{x}$ if

- L1 $z^* \in N(D, g(\bar{x}))$
- L2 $0 \in \partial f(\bar{x}) + \partial(z^* \circ g)(\bar{x}) + N(C, \bar{x})$

Let $KKT(\bar{x})$ denotes the set of KKT Lagrange multiplier for (P) at $\bar{x}$.

Calmness and metric regularity

Consider the system

$$x \in C, \quad g(x) \in D. \tag{7}$$

The system (7) is said to be *calm* (resp. *metrically regular*) at $\bar{x}$ if there exist $a > 0$ and $r > 0$ such that

$$d_{g^{-1}(D)\cap C}(x) \leq a\big(d_D(g(x)) + d_C(x)\big) \quad \forall x \in B(\bar{x}, r)$$

$\big($resp.

$$d_{g^{-1}(D-y)\cap C}(x) \leq a\big(d_D(g(x) + y) + d_C(x)\big) \quad \forall x \in B(\bar{x}, r), \forall y \in B(0, r)\big).$$

Example 1: Calmness of linear inequality systems

Consider the linear inequality system

$$\langle x_i^*, x\rangle + b_i \leq 0 \quad i = 1, \cdots, m$$

with solution set S, $b_i \in R$ and $x_i^* \in H$, with $\|x_i^*\| = 1$, where H is a Hilbert space.

A.J.

Two following properties hold and are equivalent
(i) there exists $\alpha > 0$ depending only on $(x_i^*)_{i\in\Delta_m}$ such that

$$d_S(x) \leq \alpha f(x), \quad \forall x \in X$$

(ii) (*Farkas Lemma*) for all u in S, $N(S,u) = R_+\partial f(u)$
where $f(x) = \sum_{i=1}^{m} \max(0, \langle x_i^*, x\rangle + b_i)$ and $N(S,u) = R_+\partial d(u,S)$.

Example 2: Calmness of linear equality-inequality systems

the following linear equality-inequality systems

$$Ax = 0, \quad \langle x_i^*, x\rangle + b_i \leq 0 \quad i = 1, \cdots, m \tag{8}$$

with solution set S. Here $A : X \to Y$ is a linear continuous mapping such that $R(A)$, the rang of A, is closed, $b_i \in R$ and $x_i^* \in X^*$, with $\|x_i^*\| = 1$.

Ioffe, A.J.

Then there exists $a > 0$ such that

$$d_S(x) \leq af(x), \quad \forall x \in X$$

where $f(x) = \|Ax\| + \sum_{i=1}^{m} (\langle x_i^*, x\rangle + b_i)_+$.

Example 3: Necessary and/or sufficient conditions for calmness of convex systems

Burke and Ferris

Suppose X is a Hilbert space and f is convex and proper. If S is closed, then the following are equivalent

(i) There exists $a > 0$ such that

$$d_S(x) \leq af_+(x) \quad \forall x \in X$$

(ii) $\partial d(S, x) \subset a\partial f_+(x)$ for all $x \in S$.

Where $f_+(x) = \max(f(x), 0)$

Under Slater condition

Suppose f convex and $f(x_0) < 0$. Then for all $x \in X$

$$d_S(x) \leq \frac{f_+(x)}{-f(x_0)} \|x - x_0\|.$$

Example 4: Calmness of eigenvalue matrix inequality systems

The eigenvalues of the symetric matrix X are $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$.

J. Ye and A.J.

For any given constants α_i, c such that $\sum_{i=1}^n \alpha_i \neq 0$, the set

$$S_1 := \{X \in \mathcal{S}^n : \sum_{i=1}^n \alpha_i \lambda_i(X) \leq c\}$$

is nonempty and

$$d(X, S_1) \leq \frac{\sqrt{n}}{|\sum_{i=1}^n \alpha_i|}(\sum_{i=1}^n \alpha_i \lambda_i(X) - c), \quad \forall X \notin S_1. \tag{9}$$

If moreover $\alpha_1 \geq \cdots \geq \alpha_n > 0$ and $c = 0$ or $\alpha_1 = \alpha_2 = \cdots = \alpha_n > 0$ then $\frac{\sqrt{n}}{\sum_{i=1}^n \alpha_i}$ is the smallest constant for which inequality (9) holds.

Example 5: Calmness of eigenvalue matrix inequality systems

For an integer κ between 1 and n, consider the function

$$E_\kappa(X) := \text{ sum of the } \kappa\text{th largest eigenvalues of } X.$$

Then it is clear that

$$E_\kappa(X) = \sum_{i=1}^{\kappa} \lambda_i(X) = \sum_{i=1}^{n} \alpha_i \lambda_i(X) \quad \forall X \in \S^n,$$

with $\alpha_i = 1,\ i = 1, \ldots, \kappa$ and $\alpha_i = 0,\ i = \kappa + 1, \ldots, n$.

D. Azé and J.-B. Hiriart-Urruty

Let $S_4 := \{X : E_\kappa(X) \leq c\}$. Then the set S_4 is nonempty and

$$d(X, S_4) \leq \frac{\sqrt{n}}{\kappa} \left(E_\kappa(X) - c\right), \quad \forall X \notin S_4.$$

Moreover, if either $c = 0$ or $\kappa = n$, then the constant $\frac{\sqrt{n}}{\kappa}$ is the smallest one satisfying the last inequality.

Example 6: Calmness of eigenvalue matrix inequality systems

For integers k, l between 1 and n, with $k \leq l$, consider the function

$$KL(X) := \text{sum of the } k\text{th and } l\text{th largest eigenvalues of } X.$$

Then it is clear that

$$KL(X) = \lambda_k(X) + \lambda_l(X) = \sum_{i=1}^{n} \alpha_i \lambda_i(X) \quad \forall X \in \S^n,$$

with $\alpha_i = 1,\ i = k, l$ and $\alpha_i = 0,\ i \neq k$ or $i \neq l$.

J. Ye and A. J.

Let $S_6 := \{X : KL(X) \leq c\}$. Then S_6 is nonempty and

$$d(X, S_6) \leq \frac{\sqrt{n}}{s(k,l)}(KL(X) - c), \quad \forall X \notin S_6,$$

where $s(k,l) = 1$ if $k = l$ and $s(k,l) = 2$ if $k \neq l$.

Example 7: Sufficient conditions for calmness of nonconvex inequality systems

The positive linear independence condition: $C = X$, $D = \mathbb{R}^m_+$

$$\sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x}) = 0, \quad \lambda_i \geq 0, i = 1, \cdots, m \implies \lambda_1 = \cdots = \lambda_m = 0.$$

Robinson condition : C and D are convex

$$0 \in \operatorname{int}\big(Dg(\bar{x})(C - \bar{x}) - (D - g(\bar{x}))\big)$$

or equivalently

$$Dg(\bar{x})\big(T(C, \bar{x})\big) - T(D, g(\bar{x})) = Z.$$

Rockafellar condition : $D \subset \mathbb{R}^m$

$$z^* \in N(D, g(\bar{x})), \quad 0 \in \partial(z^* \circ g)(\bar{x}) + N(C, \bar{x}) \implies z^* = 0$$

KKT multipliers

Let $\bar{x}$ be a solution of the problem (P) at which f and g are locally Lipschitz. Suppose that the system (7) is calm at $\bar{x}$. Then there exists $y^* \in N(D, g(\bar{x}))$ such that

$$0 \in \partial f(\bar{x}) + \partial(y^* \circ g)(\bar{x}) + N(C, \bar{x}).$$

The generality may exclude simple cases

Refinements of necessary optimality conditions : $D = \{0\}$

$$\begin{cases} \min f(x) \\ s.c. \quad g(x) = 0, \quad x \in C \end{cases} \tag{10}$$

(H_f) f is Gâteaux differentiable at x_0 and locally Lipschitz around x_0 with constant $K_f > 0$

(H_g^1) g is continuous and Gâteaux differentiable around x_0 .

We consider the following system

$$\text{Find} \quad x \in C, \quad g(x) = 0 \tag{11}$$

Theorem

Let x_0 be a local solution of the problem (10) where the system (11) is calm. Suppose that (H_f) and (H_g^1) hold. Then there exists $y^* \in Y^*$ such that

$$-Df(x_0) - D^*g(x_0)y^* \in N(C, x_0).$$

Refinements of necessary optimality conditions : $D \subsetneq Y$ closed set

$$\begin{cases} \min f(x) \\ s.c. \quad g(x) \in D, \quad x \in C \end{cases} \tag{12}$$

(H_f) f is Gâteaux differentiable at x_0 and locally Lipschitz around x_0.

(H_g^2) g is Gâteaux differentiable at x_0 and locally Lipschitzian around x_0.

Theorem

Let x_0 be a local solution of the problem (12) where the system (7) is calm. Suppose that (H_f) and (H_g^2) hold. Then there exists $y^* \in N(D, g(x_0))$ such that

$$-Df(x_0) - D^*g(x_0)y^* \in N(C, x_0).$$

Connection with the subdifferential of the value function

To the problem (P), we associate the family of problems

$$(P(y)) \quad \begin{cases} \min f(x) \\ x \in C \\ g(x) + y \in D \end{cases}$$

Let $v(y)$ be the value of this problem, that is, $v(y) := \inf(P(y))$ and $S(y)$ be the solution set of $(P(y))$.

Estimating the subdifferential of the value function

Suppose that $\dim F < \infty$, f and g are locally Lipschitz at any $\bar{x} \in S(0)$, and there exists a compact set K such that

$$S(y) \subset K \quad \text{for y near 0.}$$

Suppose also that the system (7) is metrically regular at any $\bar{x} \in S(0)$. Then v is locally Lipschitz at 0 and

$$\partial_\ell v(0) \subset \bigcup_{\bar{x} \in S(0)} KKT(\bar{x}).$$

Subdifferential of the value function

Let us consider the following family of problems where the data depend on the parameter

$$(Q(y)) \quad \begin{cases} \min f(x,y) \\ x \in C \\ g(x,y) \in D \end{cases}$$

Let $v(y)$ be the value of this problem, that is, $v(y) := \inf(Q(y))$ and $S(y)$ be the solution set of $(Q(y))$.

Partial metric regularity

We say that the system

$$x \in C, \quad g(x,y) \in D \tag{13}$$

is partially metrically regular at $\bar{x}$ with respect to $\bar{y}$ if there exist $a > 0$ and $r > 0$ such that

$$d_{g_y^{-1}(D)\cap C}(x) \leq a\big(d_D(g(x,y)) + d_C(x)\big) \quad \forall x \in B(\bar{x}, r), \forall y \in B(\bar{y}, r),$$

where $g_y^{-1}(D) := \{x \in E : \ g(x,y) \in D\}$.

Subdifferential of the value function

Suppose that $\dim F < \infty$, f and g are locally Lipschitz at any, $(\bar{x}, 0)$, with $\bar{x} \in S(0)$, and there exists a compact set K such that

$$S(y) \subset K \quad \text{for y near 0.} \tag{14}$$

Suppose also that the system (13) is partially metrically regular at any $\bar{x} \in S(0)$ with respect to 0. Then v is locally Lipschitz at 0 and

$$\partial_\ell v(0) \subset \bigcup_{\bar{x} \in S(0)} \{y^* \in F^* : (0, y^*) \in \partial_\ell f(\bar{x}, 0) + \partial_\ell (z^* \circ g)(\bar{x}, 0) + N_\ell(C, \bar{x}) \times \{0\},\ z^* \in N_\ell(D, g(\bar{x}, 0))\}.$$

Optimal Control

The value function in optimal control

Consider the following parametrized optimal control problems

$$(Q(\tau,\omega)) \quad \begin{cases} \min f(x(T)) \\ \dot{x}(t) \in F(x(t)) \, a.e. \, t \in [\tau, T] \\ x(\tau) = \omega \end{cases} \tag{15}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a continuous function and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisfying

$$\forall x \in \mathbb{R}^n, \quad F(x) \text{ is a nonempty compact convex set} \tag{16}$$

$$F \text{ is upper semicontinuous} \tag{17}$$

$$\exists \gamma > 0,\, \beta > 0; \sup_{y \in F(x)} \|y\| \leq \gamma \|x\| + \beta,\, \forall x. \tag{18}$$

Let $v(\tau,\omega)$ be the value of this problem, that is, $v(\tau,\omega) := \inf(Q(\tau,\omega))$ and $S(\tau,\omega)$ be the solution set of $(Q(\tau,\omega))$.

Properties of v

The lower Hamiltonian h and the upper Hamiltonian corresponding to F are defined by

$$h(x,p) := \min_{v \in F(x)} \langle p, v \rangle \quad \text{and} \quad H(x,p) := \max_{v \in F(x)} \langle p, v \rangle.$$

Proposition

- $S(\tau, \omega) \neq \emptyset$.
- If F is locally Lipschitz, that is, for all $x \in \mathbb{R}^n$ there exists $K_x > 0$ and $r_x > 0$ such that

$$F(u) \subset F(v) + K_x \|u - v\| \mathbb{B} \quad \forall u, v \in B(x, r_x) \tag{19}$$

then v is continuous. Moreover v is locally Lipschitz provided f is.

Exercise

Compute the proximal and limiting proximal subdifferentials of v.

Verification function

The augmented Hamiltonian is the function $\bar{h}$ defined on $\mathbb{R}^n \times [0, T] \times \mathbb{R}^n$ by

$$\bar{h}(x, \theta, p) = \theta + h(x, p).$$

Let $\bar{x}$ be a feasible arc for $(Q(0, x_0))$.

Verification function

A continuous function $\varphi : [0, T] \times \mathbb{R}^n \mapsto \mathbb{R}$ is a verification function for $\bar{x}$ if

- $\bar{h}(x, 1, \partial_p \varphi(t, x)) \geq 0 \ \forall (t, x) \in [0, T] \times \mathbb{R}^n$,
- $\varphi(T, \cdot) = f(\cdot)$ and $\varphi(0, x_0) = f(\bar{x}(T))$.

Theorem

A feasible arc $\bar{x}$ is optimal iff there exists a continuous verification function for $\bar{x}$; the value function v is one of such verification function for any optimal arc.

The minimal time problem

The minimal time control problem consists of a given closed set S (the "target set") and a control system in which the goal is to steer an initial point ω to the target set along a trajectory of the system in minimal time. The minimal time value is denoted by $T_S(\omega)$, which could be $+\infty$ if no trajectory from ω can reach S. The system involved is governed by the differential inclusion considered in (15). So

$$T_S(\omega) = \inf\{T \geq 0 : \text{ some trajectory } x \text{ satisfies } x(0) = \omega,\ x(T) \in S\}.$$

Theorem

Suppose F satisfies (16)-(19). Then there exists a unique lower semicontinuous function $\varphi : \mathbb{R}^n \mapsto]-\infty, +\infty]$ bounded below on $\mathbb{R}^n$ and satisfying the following:

- $\forall x \notin S, \quad h(x, \partial_p \varphi(x)) = -1;$
- Each $x \in S$ satisfies $\varphi(x) = 0$ and $h(x, \partial_p \varphi(x)) \geq -1.$

The minimal time problem

The minimal time control problem consists of a given closed set S (the "target set") and a control system in which the goal is to steer an initial point ω to the target set along a trajectory of the system in minimal time. The minimal time value is denoted by $T_S(\omega)$, which could be $+\infty$ if no trajectory from ω can reach S. The system involved is governed by the differential inclusion considered in (15). So

$$T_S(\omega) = \inf\{T \geq 0 : \text{ some trajectory } x \text{ satisfies } x(0) = \omega,\ x(T) \in S\}.$$

Theorem

Suppose F satisfies (16)-(19). Then there exists a unique lower semicontinuous function $\varphi : \mathbb{R}^n \mapsto]-\infty, +\infty]$ bounded below on $\mathbb{R}^n$ and satisfying the following:

- $\forall x \notin S, \quad h(x, \partial_p \varphi(x)) = -1;$
- Each $x \in S$ satisfies $\varphi(x) = 0$ and $h(x, \partial_p \varphi(x)) \geq -1.$

The unique such function is $\varphi(\cdot) = T_S(\cdot)$.

Violation of the lower semicontinuity of T_S and the existence of optimal trajectory

The following examples show the necessity of the assumption (18). Define $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ by

$$F(x, y) = \{(1, 1 + y^2)\}.$$

Violation of the lower semicontinuity of T_S

Consider the target set $S = \{\frac{\pi}{2}\} \times \mathbb{R}$. Then one has $T_S(0, 0) = \infty$, while $\lim_{s \to 0^+} T_S(s, 0) = \frac{\pi}{2}$.

Violation of the existence of optimal trajectory

Consider the target set $S = \{(x,y) \in \mathbb{R}^2 :\ x \geq \frac{\pi}{2},\ y(x - \frac{\pi}{2}) = 1\}$. Then the reachable set $R^T(0,0)$ from $(0,0)$ at time T, that is,

$$R^T(0,0) := \{x(T) \in S :\ x \text{ is a trajectory for } F, x(0) = (0,0)\},$$

is given by

$$R^T(0,0) = \begin{cases} \{T\} \times [0, \tan T] & \text{if } 0 \leq T < \frac{\pi}{2} \\ \{T\} \times [0, +\infty[& \text{if } T \geq \frac{\pi}{2}. \end{cases}$$

Thus as

$$T_S(0,0) = \inf\{T \geq 0 :\ R^T(0,0) \cap S \neq \emptyset\}$$

then one has $T_S(0,0) = \frac{\pi}{2}$ but no trajectory reaches S from $(0,0)$ in this time.

The proximal subdifferential of the minimal time function

For $r \geq 0$, define

$$S(r) := \{\omega \in \mathbb{R}^n : \; T_S(\omega) \leq r\}$$

the $r-$level set of $T_S(\cdot)$.

Computation of the proximal subdifferential

Suppose F satisfies (16)-(19). Then

- For all $x \in S$, we have

$$\partial_p T_S(x) = N_p(S, x) \cap \{p \in \mathbb{R}^n : \; h(x, p) \geq -1\}.$$

- Whenever $r > 0$ and $T_S(x) = r$, then

$$\partial_p T_S(x) = N_p(S(r), x) \cap \{p \in \mathbb{R}^n : \; h(x, p) = -1\}.$$

Lipschitz continuity of the minimal time function

Characterization of the Lipschitz continuity

Suppose F satisfies (16)-(19). Then the following are equivalent:

- There exists $\eta > 0$ such that $T_S(\cdot)$ is Lipschitz continuous on $S + \eta B$.
- $$\sup_{x \in S, p \in \partial_p T_S(x)} \|p\| < \infty.$$
- There exist $\eta > 0$ and $\delta > 0$ such that
$$x \in S^c \cap (S + \eta B), \quad p \in x - P_S(x) \implies h(x, p) \leq -\delta \|p\|.$$

Necessary optimality conditions of free time problems

The free time problem are optimal control problems where the minimization is given jointly in time and state:

$$(FT(\omega)) \quad \begin{cases} \min\limits_{(T,x)} f(T, x(T)) \\ \dot{x}(t) \in F(x(t)) \; a.e. \; t \in [0, T] \\ x(0) = \omega \end{cases} \tag{20}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a locally Lipschitzian function and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisfying (16)-(19).
Let $v(\omega)$ be the value of this problem and $\bar{T} > 0$.

Necessary optimality conditions

Let $(\bar{T}, \bar{x})$ be a solution to the problem (20). Then there exists an arc p on $[0, \bar{T}]$ such that

- $(-\dot{p}(t), \dot{\bar{x}}(t)) \in \partial_c H(\bar{x}(t), p(t))$, a.e. $t \in [0, \bar{T}]$;
- $H(\bar{x}(t), p(t)) = a\,(=$ constant$)$, $0 \leq t \leq \bar{T}$ and $(a, -p(\bar{T})) \in \partial_\ell f(\bar{T}, \bar{x}(\bar{T}))$.

Invariance

Let $S \subset \mathbb{R}^n$ be a closed set.

Weak invariance

The system (S, F) is said to be weakly invariant provided that for all $x_0 \in S$, there exists a trajectory x on $[0, +\infty[$ such that

$$x(0) = x_0, \quad x(t) \in S \, \forall t \geq 0.$$

Theorem

Let F satisfying (16)-(18). Then the following are equivalent:

- (S, F) is weakly invariant;
- $h(x, N_p(S, x)) \leq 0 \ \forall x \in S$;
- $F(x) \cap K(S, x) \neq \emptyset \ \forall x \in S$;
- $F(x) \cap \mathrm{co}K(S, x) \neq \emptyset \ \forall x \in S$.

Here $K(S, x)$ denotes the contingent or Bouligand cone to S at x.

Characterization of solutions of Hamilton-Jacobi equations

Theorem

A continuous function u on $\Omega \subset \mathbb{R}^n$ is

- a viscosity supersolution of (6) iff for all $x \in \Omega$

$$H(x, u(x), x^*) \geq 0 \quad \forall x^* \in \partial_F u(x)$$

- a viscosity subsolution of (6) iff for all $x \in \Omega$

$$H(x, u(x), x^*) \leq 0 \quad \forall x^* \in \partial_F^+ u(x).$$

Here $\partial_F^+ u(x) = -\partial_F(-u)(x)$.

Characterization of test functions

Lemma

Let u be a continous function on Ω. Then

(i) $x^* \in \partial_F u(x)$ if and only if there exists a function $\varphi \in \mathcal{C}^1(\Omega)$ such that $\nabla\varphi(x) = x^*$ and $u - \varphi$ has a local minimum at x.

(ii) $x^* \in \partial_F^+ u(x)$ if and only if there exists a function $\varphi \in \mathcal{C}^1(\Omega)$ such that $\nabla\varphi(x) = x^*$ and $u - \varphi$ has a local maximum at x.

Example

Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \mapsto \mathbb{R}$ be a continuous function which is $1-$Lipschtit on $\partial\Omega$. Then the function $\varphi : \Omega \mapsto \mathbb{R}$ defined by $\varphi(x) = \inf_{y \in \partial\Omega} \{\|y - x\| + f(y)\}$ is a viscosity subsolution of the Hamilton-Jacobi equation

$$\begin{cases} |\nabla\varphi(x)| = 1 & x \in \Omega \\ \varphi_{\partial\Omega} = f \end{cases}$$

- V. Bompart, Optimisation non lisse pour la commande des systèmes de l'Aéronautique, Doctorat de l'Université Paul Sabatier, Toulouse III, 2007.
- P. Bosch, R. Henrion, A. Jourani, Error bounds and applications, Applied Math. Optim., 50 (2004) 161-181.
- F.H. Clarke, Optimisation and nonsmooth analysis, Wiley, New-York (1983).
- F.H. Clarke, Necessary conditions in dynamic programming, *Memoirs AMS* **73** (2005).
- F. H. Clarke, Discontinuous Feedback and Nonlinear Systems, Proc. IFAC Conf. Nonlinear Control (NOLCOS), Bologna, 2010, p.1-29.
- F.H. Clarke, Yu. S. Ledyaev, R.J. Stern and P. R. Wolenski, Nonsmooth Analysis and Control Theory, Graduate Texts in Mathematics, Springer, 1998.

- M. Fabian, Subdifferentiability and trustworthiness in the light of new variational principle of Borwein and Preiss, Acta Univ. Carolinae (1989), 51-56.
- H. Frankowska, Local controllability and infinitesimal generators of semigroups of set-valued maps, SIAM J. Cont. Optim., 25 (1987) 412–432.
- H. Frankowska, B. Kaskosz, A Maximum principle for differential inclusion problems with state constraints, Systems Control Lett., 11 (1988) 189-194.
- A. Jourani, L. Thibault, Metric regularity and subdifferential calculus in Banach spaces, Set-valued Anal., 3 (1995), 87-100.
- A. Jourani, L. Thibault, Verifiable conditions for openness and metric regularity of multivalued mappings in Banach spaces, Transactions AMS, 347 (1995), 1255-1268.

A. Jourani, L. Thibault, Noncoincidence of approximate and limiting subdifferentials of integral functionals, SIAM J. Cont. Optim., 49 (2011), 1435-1453

B.S. Mordukhovich, Variational Analysis and Generalized Differentiation, Vol.1: Basic Theory, Vol. 2: Applications, Springer, Berlin (2005).

P. Wolenski and Y. Zhuang, Proximal analysis and the minimal time function, SIAM J. Cont. Optim.,36 (1998), 1048-1072.

Luděk Zajíček, Strict differentiability via differentiability Acta Universitatis Carolinae. Mathematica et Physica, Vol. 28 (1987), 157-159