The gradient's limit of a definable family of functions is a conservative field

Sholom Schechtman

Introduction

Let $(f_a)_{a>0}$ be a family of smooth functions and $F : \mathbb{R}^d \to \mathbb{R}$ be smooth such that

$$\|f_a - F\| \xrightarrow[a \to 0]{} 0\,.$$

Does the following holds?

$$\|\nabla f_a - \nabla F\| \xrightarrow[a \to 0]{?} 0$$

Introduction

More formally,

$$\|f_a - F\| \xrightarrow[a \to 0]{} 0\,.$$

Introduction

More formally,

$$\|f_a - F\| \xrightarrow[a\to 0]{} 0\,.$$

Denote the gradients limits:

$$D_F(x) = \{v \in \mathbb{R}^d : x_n \to x\,, a_n \to 0\,, \nabla f_{a_n}(x_n) \to v\}\,.$$

Introduction

More formally,

$$\|f_a - F\| \xrightarrow[a \to 0]{} 0\,.$$

Denote the gradients limits:

$$D_F(x) = \{v \in \mathbb{R}^d : x_n \to x\,, a_n \to 0\,, \nabla f_{a_n}(x_n) \to v\}\,.$$

Does the following holds?

$$D_F(x) = \{\nabla F(x)\}\,.$$

Introduction

More generally

$$f_a \xrightarrow[a\to 0]{} F, \text{ with } f_a : \mathbb{R}^d \to \mathbb{R} \text{ locally Lipschitz.}$$

Introduction

More generally

$$f_a \xrightarrow[a \to 0]{} F, \text{ with } f_a : \mathbb{R}^d \to \mathbb{R} \text{ locally Lipschitz.}$$

$$D_F(x) = \left\{ v \in \mathbb{R}^d : \text{there is } (x_n, v_n, a_n) \to (x, v, 0) \text{ with } v_n \in \partial f_{a_n}(x_n) \right\}.$$

➔ $\partial f_a(x)$ some subgradient (convex, Frechet, Clarke) of f_a at x.

Introduction

More generally

$$f_a \xrightarrow[a\to 0]{} F, \text{ with } f_a : \mathbb{R}^d \to \mathbb{R} \text{ locally Lipschitz.}$$

$$D_F(x) = \{v \in \mathbb{R}^d : \text{there is } (x_n, v_n, a_n) \to (x, v, 0) \text{ with } v_n \in \partial f_{a_n}(x_n)\}\,.$$

➔ $\partial f_a(x)$ some subgradient (convex, Frechet, Clarke) of f_a at x.

Question. What is the link between D_F and ∂F?

Introduction

More generally

$$f_a \xrightarrow[a \to 0]{} F, \text{ with } f_a : \mathbb{R}^d \to \mathbb{R} \text{ locally Lipschitz.}$$

$$D_F(x) = \{v \in \mathbb{R}^d : \text{there is } (x_n, v_n, a_n) \to (x, v, 0) \text{ with } v_n \in \partial f_{a_n}(x_n)\}\,.$$

➔ $\partial f_a(x)$ some subgradient (convex, Frechet, Clarke) of f_a at x.

Question. What is the link between D_F and ∂F?

- One would wish that D_F reduces to some common first-order operators.
  - ➔ To the gradient if F is smooth.
  - ➔ To the subgradient if F is not.

Introduction

$f_a \to F$

$$D_F(x) = \{v \in \mathbb{R}^d : \text{there is } (x_n, v_n, a_n) \to (x, v, 0) \text{ with } x_n \in \partial f_a(x_n)\}\,.$$

Question. Link between D_F and ∂F?

Motivation.

Introduction

$f_a \to F$

$D_F(x) = \{v \in \mathbb{R}^d : \text{there is } (x_n, v_n, a_n) \to (x, v, 0) \text{ with } x_n \in \partial f_a(x_n)\}\,.$

Question. Link between D_F and ∂F?

Motivation.

1) Interesting in its own right.

Introduction

$f_a \to F$

$$D_F(x) = \{v \in \mathbb{R}^d : \text{there is } (x_n, v_n, a_n) \to (x, v, 0) \text{ with } x_n \in \partial f_a(x_n)\}\,.$$

Question. Link between D_F and ∂F?

Motivation.

1) Interesting in its own right.
2) Consequences for smoothing methods.

Smoothing methods

$$\text{Optimization problem:} \quad \min_{x \in \mathbb{R}^d} F(x)\,.$$

- Might be hard if F is nonsmooth and nonconvex.

Smoothing methods

Optimization problem: $\min_{x \in \mathbb{R}^d} F(x)\,.$

- Might be hard if F is nonsmooth and nonconvex.
- Construct $f_a \to F$, with f_a smooth.

Smoothing methods:

Smoothing methods

$$\text{Optimization problem:} \quad \min_{x \in \mathbb{R}^d} F(x)\,.$$

- Might be hard if F is nonsmooth and nonconvex.
- Construct $f_a \to F$, with f_a smooth.

Smoothing methods:

- For $a_k, \varepsilon_k > 0$,

$$\text{Solve} \quad \min_{x \in \mathbb{R}^d} f_{a_k}(x)$$

Smoothing methods

Optimization problem: $$\min_{x \in \mathbb{R}^d} F(x) \,.$$

- Might be hard if F is nonsmooth and nonconvex.
- Construct $f_a \to F$, with f_a smooth.

Smoothing methods:

- For $a_k, \varepsilon_k > 0$,

$$\text{Solve} \quad \min_{x \in \mathbb{R}^d} f_{a_k}(x) \quad \Longrightarrow \text{ find } \|\nabla f_{a_k}(x_k)\| \leq \varepsilon_k \,.$$

Smoothing methods

Optimization problem: $\min_{x \in \mathbb{R}^d} F(x)\,.$

- Might be hard if F is nonsmooth and nonconvex.
- Construct $f_a \to F$, with f_a smooth.

Smoothing methods:

- For $a_k, \varepsilon_k > 0$,

$$\text{Solve} \quad \min_{x \in \mathbb{R}^d} f_{a_k}(x) \quad \Longrightarrow \text{ find } \|\nabla f_{a_k}(x_k)\| \leq \varepsilon_k\,.$$

➔ Easier because f_{a_k} is smooth.

Smoothing methods

Optimization problem: $\min_{x \in \mathbb{R}^d} F(x)\,.$

- Might be hard if F is nonsmooth and nonconvex.
- Construct $f_a \to F$, with f_a smooth.

Smoothing methods:

- For $a_k, \varepsilon_k > 0$,

$$\text{Solve} \quad \min_{x \in \mathbb{R}^d} f_{a_k}(x) \quad \Longrightarrow \text{find } \|\nabla f_{a_k}(x_k)\| \leq \varepsilon_k\,.$$

  - ➔ Easier because f_{a_k} is smooth.
- Decrease a_k and ε_k and repeat.

Smoothing methods

Optimization problem: $\min_{x \in \mathbb{R}^d} F(x)\,.$

- Might be hard if F is nonsmooth and nonconvex.
- Construct $f_a \to F$, with f_a smooth.

Smoothing methods:

- For $a_k, \varepsilon_k > 0$,

$$\text{Solve} \quad \min_{x \in \mathbb{R}^d} f_{a_k}(x) \quad \Longrightarrow \text{ find } \|\nabla f_{a_k}(x_k)\| \leq \varepsilon_k\,.$$

  ➔ Easier because f_{a_k} is smooth.

- Decrease a_k and ε_k and repeat.
- When $a_k, \varepsilon_k \to 0$ and $x_k \to x^*$

$$\lim_{k \to +\infty} \nabla f_{a_k}(x_k) = 0$$

Smoothing methods

Optimization problem: $\min_{x \in \mathbb{R}^d} F(x)$.

- Might be hard if F is nonsmooth and nonconvex.
- Construct $f_a \to F$, with f_a smooth.

Smoothing methods:

- For $a_k, \varepsilon_k > 0$,

$$\text{Solve} \quad \min_{x \in \mathbb{R}^d} f_{a_k}(x) \quad \Longrightarrow \text{ find } \|\nabla f_{a_k}(x_k)\| \leq \varepsilon_k \,.$$

  ➔ Easier because f_{a_k} is smooth.
- Decrease a_k and ε_k and repeat.
- When $a_k, \varepsilon_k \to 0$ and $x_k \to x^*$

$$\lim_{k \to +\infty} \nabla f_{a_k}(x_k) = 0 \in D_F(x^*) \,.$$

Smoothing methods

Optimization problem: $\min_{x \in \mathbb{R}^d} F(x)\,.$

- Might be hard if F is nonsmooth and nonconvex.
- Construct $f_a \to F$, with f_a smooth.

Smoothing methods:

- For $a_k, \varepsilon_k > 0$,

$$\text{Solve} \quad \min_{x \in \mathbb{R}^d} f_{a_k}(x) \quad \Longrightarrow \quad \text{find } \|\nabla f_{a_k}(x_k)\| \le \varepsilon_k\,.$$

  ➔ Easier because f_{a_k} is smooth.

- Decrease a_k and ε_k and repeat.
- When $a_k, \varepsilon_k \to 0$ and $x_k \to x^*$

$$\lim_{k \to +\infty} \nabla f_{a_k}(x_k) = 0 \in D_F(x^*)\,.$$

  ➔ Necessary optimality condition if $D_F = \nabla F$ or ∂F.

Known results

Attouch's theorem

If for every $a > 0$, f_a is convex. Then $D_F(x) = \partial F(x)$.

Attouch, 1977

- With $\partial F(x)$ being convex subgradient of F at x.
- Several extensions to "approximately convex" case.
  - ➔ Poliquin, 1992; Levy, Poliquin, and Thibault, 1995; Zolezzi, 1985; Zolezzi, 1994; Czarnecki and Rifford, 2006 ...

Nonconvex case?

Can we do more?
Not really.

Nonconvex case?

Can we do more?

Not really.

- ➔ $f_a(x) = a\sin(\frac{x}{a})$.
- ➔ $\sin(\frac{x}{a}) \to [-1, 1]$.

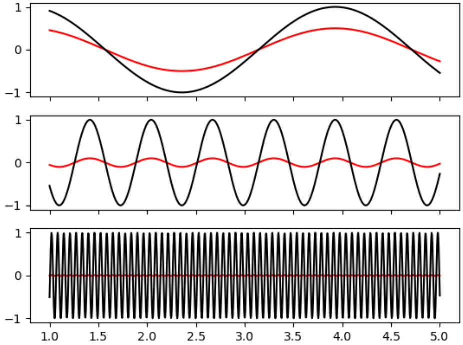


Figure: $f_a(x)$ and $f_a'(x)$

Nonconvex case?

Main result

- Assume that for each $a > 0$, f_a is locally Lipschitz continuous [1]
- $f_a \to F$ uniformly on compact sets.

Theorem

If the family (f_a) is definable, then D_F is a conservative set-valued field of F (in the sense of Bolte and Pauwels).

Schechtman, 2024

[1] Can be relaxed.

Main result

- Assume that for each $a > 0$, f_a is locally Lipschitz continuous [1]
- $f_a \to F$ uniformly on compact sets.

Theorem

If the family (f_a) is definable, then D_F is a conservative set-valued field of F (in the sense of Bolte and Pauwels).

Schechtman, 2024

Some consequences.

- For almost every $x \in \mathbb{R}^d$, $D_F(x) = \{\nabla F(x)\}$.

[1] Can be relaxed.

Main result

- Assume that for each $a > 0$, f_a is locally Lipschitz continuous [1]
- $f_a \to F$ uniformly on compact sets.

Theorem

If the family (f_a) is definable, then D_F is a conservative set-valued field of F (in the sense of Bolte and Pauwels).

Schechtman, 2024

Some consequences.

- For almost every $x \in \mathbb{R}^d$, $D_F(x) = \{\nabla F(x)\}$.
- $\partial F(x) \subset \operatorname{conv} D_F(x)$
  - ➔ $0 \in \operatorname{conv} D_F(x)$ is a necessary optimality condition.

[1] Can be relaxed.

Main result

- Assume that for each $a > 0$, f_a is locally Lipschitz continuous [1]
- $f_a \to F$ uniformly on compact sets.

Theorem

If the family (f_a) is definable, then D_F is a conservative set-valued field of F (in the sense of Bolte and Pauwels).

Schechtman, 2024

Some consequences.

- For almost every $x \in \mathbb{R}^d$, $D_F(x) = \{\nabla F(x)\}$.
- $\partial F(x) \subset \operatorname{conv} D_F(x)$
  - ➔ $0 \in \operatorname{conv} D_F(x)$ is a necessary optimality condition.
- The geometric structure of D_F is well understood.

[1] Can be relaxed.

Outline

I Definable/tame functions

II Conservative set-valued fields

III Convergence of gradients

Outline

I Definable/tame functions
II Conservative set-valued fields
III Convergence of gradients

Outline

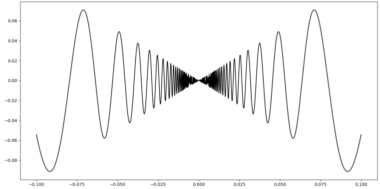
I Definable/tame functions
II Conservative set-valued fields
III Convergence of gradients

Definable functions

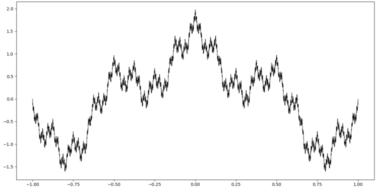
A continuous function in $\mathbb{R}^d$ can be pathological.

Definable functions

A continuous function in $\mathbb{R}^d$ can be pathological.



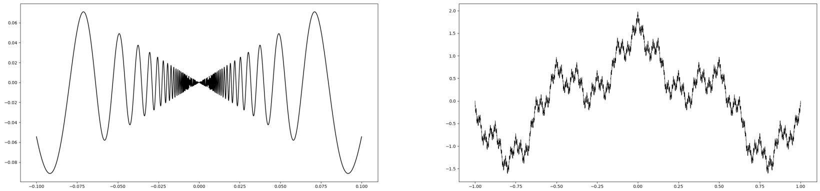
(a) $x \mapsto x \sin(1/x)$



(b) Weierstrass function

Definable functions

A continuous function in $\mathbb{R}^d$ can be pathological.



(a) $x \mapsto x \sin(1/x)$

(b) Weierstrass function

There are functions $F : [0,1] \to [0,1]$ such that $\forall x\,,\ \partial F(x) = [0,1]$

Definable functions

A class of function that includes:

- Semialgebraic: defined by a finite number of polynomial inequalities.

Definable functions

A class of function that includes:

- Semialgebraic: defined by a finite number of polynomial inequalities.
  - ➔ Piecewise polynomials.

Definable functions

A class of function that includes:

- Semialgebraic: defined by a finite number of polynomial inequalities.
  - ➔ Piecewise polynomials.
  - ➔ Piecewise rational functions.

Definable functions

A class of function that includes:

- Semialgebraic: defined by a finite number of polynomial inequalities.
  - ➔ Piecewise polynomials.
  - ➔ Piecewise rational functions.
  - ➔ $x \mapsto \|x\|_q$, with $q \geq 0$ rational.

Definable functions

A class of function that includes:

- Semialgebraic: defined by a finite number of polynomial inequalities.
  - Piecewise polynomials.
  - Piecewise rational functions.
  - $x \mapsto \|x\|_q$, with $q \geq 0$ rational.
  - $\max, \min$ of semi-algebraic functions.

Definable functions

A class of function that includes:

- Semialgebraic: defined by a finite number of polynomial inequalities.
  - Piecewise polynomials.
  - Piecewise rational functions.
  - $x \mapsto \|x\|_q$, with $q \geq 0$ rational.
  - $\max, \min$ of semi-algebraic functions.
- Analytic restricted to a (semialgebraic) compact.

Definable functions

A class of function that includes:

- Semialgebraic: defined by a finite number of polynomial inequalities.
  - Piecewise polynomials.
  - Piecewise rational functions.
  - $x \mapsto \|x\|_q$, with $q \geq 0$ rational.
  - $\max, \min$ of semi-algebraic functions.
- Analytic restricted to a (semialgebraic) compact.
- exp and log

Definable functions

A class of function that includes:

- Semialgebraic: defined by a finite number of polynomial inequalities.
  - ➔ Piecewise polynomials.
  - ➔ Piecewise rational functions.
  - ➔ $x \mapsto \|x\|_q$, with $q \geq 0$ rational.
  - ➔ $\max, \min$ of semi-algebraic functions.
- Analytic restricted to a (semialgebraic) compact.
- exp and log
- Stable by $\{+, -, \times, /, \circ^{-1}\}$ and composition.

Definable functions

A class of function that includes:

- Semialgebraic: defined by a finite number of polynomial inequalities.
  - Piecewise polynomials.
  - Piecewise rational functions.
  - $x \mapsto \|x\|_q$, with $q \geq 0$ rational.
  - $\max, \min$ of semi-algebraic functions.
- Analytic restricted to a (semialgebraic) compact.
- exp and log
- Stable by $\{+, -, \times, /, \circ^{-1}\}$ and composition.

Definable functions are piecewise smooth 1/2

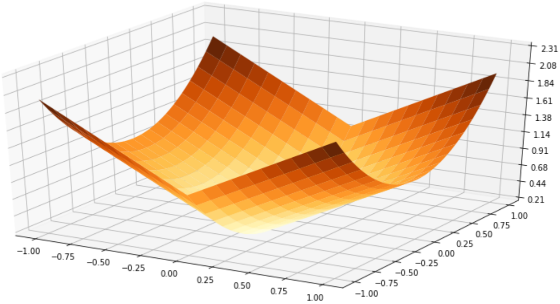


Figure: $F(y, z) = \frac{1}{2}y^2 + |z|$

- F is smooth on $X_1 = \{(y, z) : z > 0\}$ and $X_2 = \{(y, z) : z < 0\}$.

Definable functions are piecewise smooth 1/2

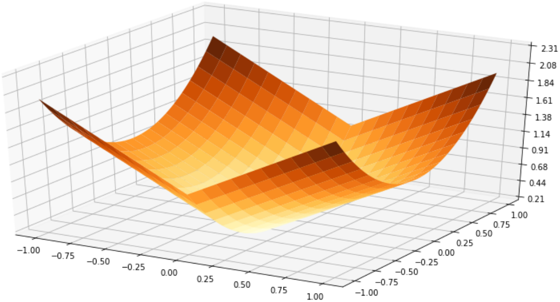


Figure: $F(y,z) = \frac{1}{2}y^2 + |z|$

- F is smooth on $X_1 = \{(y,z) : z>0\}$ and $X_2 = \{(y,z) : z<0\}$.
- F is smooth on $X_3 = \{(y,z) : z=0\}$.

Definable functions are piecewise smooth 1/2

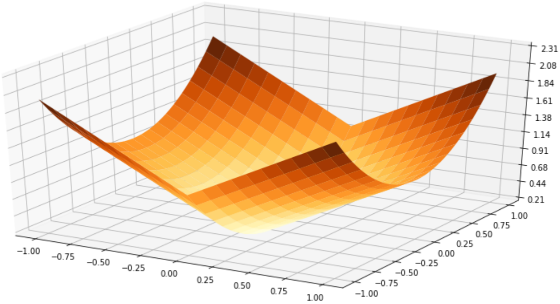


Figure: $F(y,z) = \frac{1}{2}y^2 + |z|$

- F is smooth on $X_1 = \{(y,z) : z>0\}$ and $X_2 = \{(y,z) : z<0\}$.
- F is smooth on $X_3 = \{(y,z) : z=0\}$.

Definable functions are piecewise smooth 2/2

Theorem

Let $F : \mathbb{R}^d \to \mathbb{R}$ be definable and $p \in \mathbb{N}$. There exists (X_i), a finite partition of $\mathbb{R}^d$ into C^p manifolds, s.t. $F_{|X_i}$ is C^p.

Dries and Miller, 1996

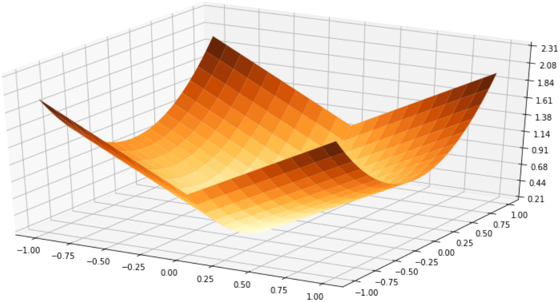


Figure: $F(y, z) = \frac{1}{2}y^2 + |z|$ is smooth on $X_1, X_2, X_3 = z > 0, z < 0, z = 0$

Main result

Theorem

If the family (f_a) is definable in an o-minimal structure, then D_F is a conservative set-valued field of F.

Schechtman, 2024

- The family (f_a) is definable if

$$(x, a) \mapsto f_a(x) \quad \text{is definable.}$$

Conservative set-valued fields

Gradient

$$F \text{ smooth} \implies \nabla F \text{ conservative}.$$

- For every a.c. curve $\mathsf{x} : [0, 1] \to \mathbb{R}^d$

Gradient

$$F \text{ smooth} \implies \nabla F \text{ conservative.}$$

- For every a.c. curve $\mathsf{x} : [0,1] \to \mathbb{R}^d$

$$F(\mathsf{x}_1) - F(\mathsf{x}_0) = \int_0^1 \langle \nabla F(\mathsf{x}_t), \dot{\mathsf{x}}_t \rangle \, \mathrm{d}t$$

$$\Updownarrow$$

$$\frac{\mathrm{d}}{\mathrm{d}t} F(\mathsf{x}_t) = \langle \nabla F(\mathsf{x}_t), \dot{\mathsf{x}}_t \rangle \, .$$

Conservative set-valued field

$F : \mathbb{R}^d \to \mathbb{R}$, locally Lipschitz continuous.

$$D_F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d \quad D(x) \subset \mathbb{R}^d \,.$$

Conservative set-valued field

$F : \mathbb{R}^d \to \mathbb{R}$, locally Lipschitz continuous.

$$D_F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d \quad D(x) \subset \mathbb{R}^d .$$

D_F is a conservative set-valued field of F if (Bolte and Pauwels, 2021).

- For every a.c. curve $\mathsf{x} : [0,1] \to \mathbb{R}^d$ and $\mathsf{v}_t \in D_F(\mathsf{x}_t)$

$$F(\mathsf{x}_1) - F(\mathsf{x}_0) = \int_0^1 \langle \mathsf{v}_t, \dot{\mathsf{x}}_t \rangle \, \mathrm{d}t$$

$$\Updownarrow$$

$$\frac{\mathrm{d}}{\mathrm{d}t} F(\mathsf{x}_t) = \langle \mathsf{v}_t, \dot{\mathsf{x}}_t \rangle \quad \text{almost everywhere.}$$

Examples

$$F(\mathsf{x}_1) = F(\mathsf{x}_0) + \int_0^1 \langle \mathsf{v}_t, \dot{\mathsf{x}}_t \rangle \, \mathrm{d}t \,.$$

- If F is C^1, then $D_F(x) = \{\nabla F(x)\}$ is a conservative field.

Examples

$$F(\mathsf{x}_1) = F(\mathsf{x}_0) + \int_0^1 \langle \mathsf{v}_t, \dot{\mathsf{x}}_t \rangle \, \mathrm{d}t \, .$$

- If F is C^1, then $D_F(x) = \{\nabla F(x)\}$ is a conservative field.
- If F is definable, then ∂F is a conservative field. Bolte et al., 2007

Examples

$$F(\mathsf{x}_1) = F(\mathsf{x}_0) + \int_0^1 \langle \mathsf{v}_t, \dot{\mathsf{x}}_t \rangle \,\mathrm{d}t\,.$$

- If F is C^1, then $D_F(x) = \{\nabla F(x)\}$ is a conservative field.
- If F is definable, then ∂F is a conservative field. Bolte et al., 2007
- Automatic differentiation in `TensorFlow` or `Pytorch` outputs an element of a conservative field. Bolte and Pauwels, 2021

Examples

$$F(\mathsf{x}_1) = F(\mathsf{x}_0) + \int_0^1 \langle \mathsf{v}_t, \dot{\mathsf{x}}_t \rangle \, \mathrm{d}t \,.$$

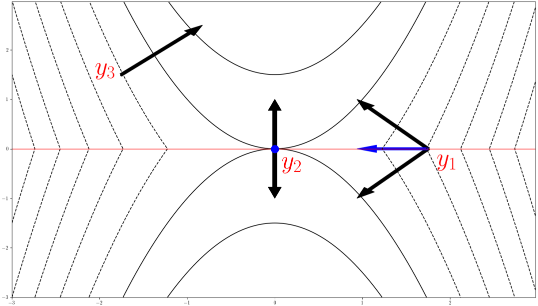
- If F is C^1, then $D_F(x) = \{\nabla F(x)\}$ is a conservative field.
- If F is definable, then ∂F is a conservative field. Bolte et al., 2007
- Automatic differentiation in `TensorFlow` or `Pytorch` outputs an element of a conservative field. Bolte and Pauwels, 2021

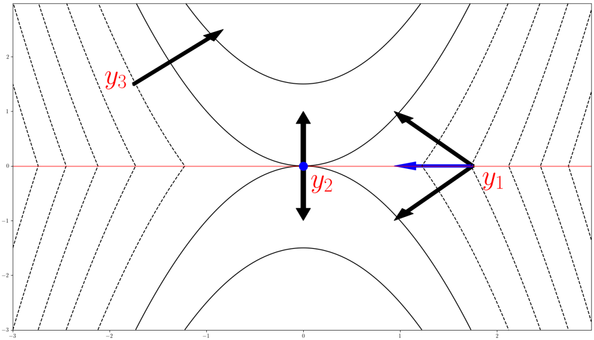
Variational stratification

Let $F : \mathbb{R}^d \to \mathbb{R}$ and $D_F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be definable. D_F is a conservative field if and only if there is a partition of $\mathbb{R}^d$ into manifolds $(\mathcal{M}_i)$ such that for every $y \in \mathcal{M}_i$

$$D_F(y) \subset \nabla_{\mathcal{M}_i} F(y) + \mathcal{N}_{\mathcal{M}_i}(y)\,.$$

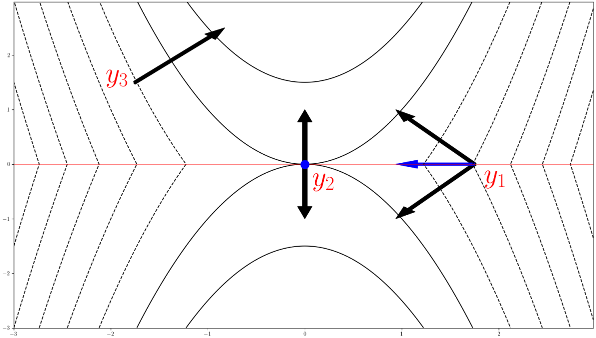
Bolte and Pauwels, 2021; Lewis and Tian, 2021; Davis and Drusvyatskiy, 2022





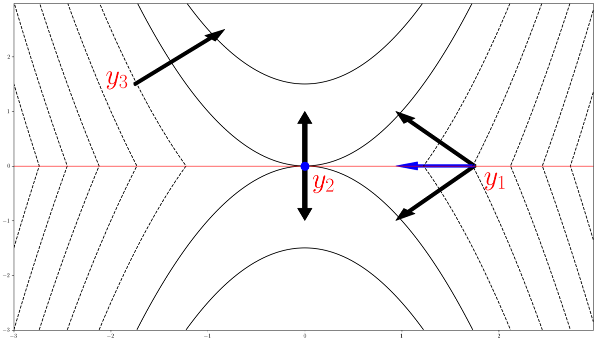
$D_F(x)$ and $\nabla_{\mathcal{M}} F(x)$

- $D_F(x) = \{\nabla F(x)\}$ on a dense open set.



$D_F(x)$ and $\nabla_{\mathcal{M}} F(x)$

- $D_F(x) = \{\nabla F(x)\}$ on a dense open set.
- ∂F is a conservative field and $\partial F(x) \subset \operatorname{conv} D_F(x)$.

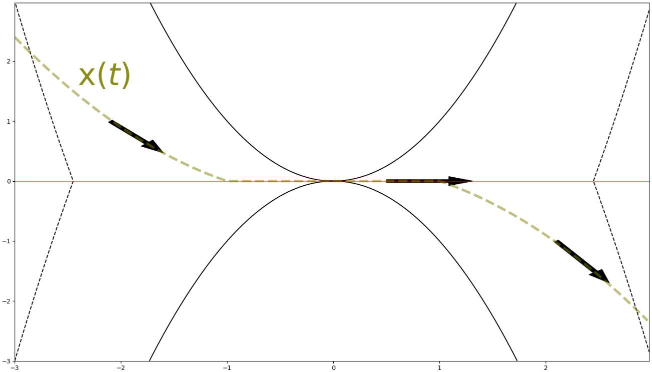


$D_F(x)$ and $\nabla_{\mathcal{M}} F(x)$

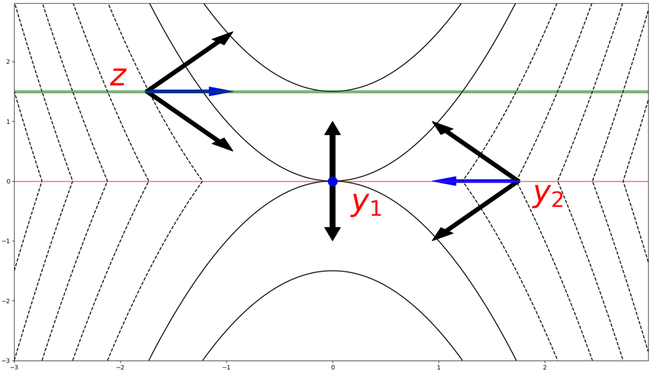
- $D_F(x) = \{\nabla F(x)\}$ on a dense open set.
- ∂F is a conservative field and $\partial F(x) \subset \operatorname{conv} D_F(x)$.

➔ x^* local minimum $\Longrightarrow 0 \in \operatorname{conv} D_F(x^*)$.

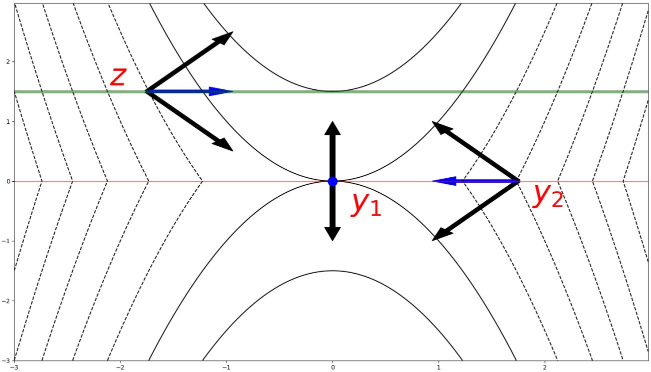
$\mathsf{x}(t)$ and $\dot{\mathsf{x}}(t)$



$$\frac{\mathrm{d}}{\mathrm{d}t}F(\mathsf{x}(t)) = \langle \mathsf{v}(t), \dot{\mathsf{x}}(t)\rangle = \langle \nabla_{\mathcal{M}} F(\mathsf{x}(t)), \dot{\mathsf{x}}(t)\rangle + \overbrace{\langle \mathcal{N}_{\mathcal{M}}(y), \dot{\mathsf{x}}(t)\rangle}^{=0} .$$

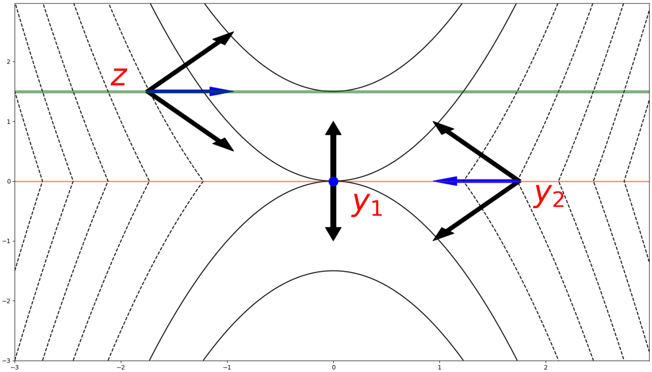


➔ Conservative fields are not unique!



➔ Conservative fields are not unique!

➔ But $D_F(x) = \{\nabla F(x)\}$ on a dense open set.



- ➔ Conservative fields are not unique!
- ➔ But $D_F(x) = \{\nabla F(x)\}$ on a dense open set.
- ➔ ∂F is a conservative field and $\partial F(x) \subset \operatorname{conv} D_F(x)$.
- ➔ x^* local minimum $\implies 0 \in \operatorname{conv} D_F(x^*)$.

Theorem

If the family (f_a) is definable in an o-minimal structure, then D_F is a conservative set-valued field of F

Schechtman, 2024

$$D_F(x) = \{v \in \mathbb{R}^d : \text{there is } (x_n, v_n, a_n) \to (x, v, 0) \text{ with } v_n \in \partial f_{a_n}(x_n)\}\,.$$

➔ ∂f_a is the Clarke's subgradient.

Convergence of gradients

Main result

- $f_a \to F$ uniformly on compact sets.

Main result

- $f_a \to F$ uniformly on compact sets.
- For each $a > 0$, D_a is a conservative field of f_a. (e.g. $D_a = \nabla f_a$ or ∂f_a).

Main result

- $f_a \to F$ uniformly on compact sets.
- For each $a > 0$, D_a is a conservative field of f_a. (e.g. $D_a = \nabla f_a$ or ∂f_a).
- The maps $(x, a) \mapsto f_a(x)$ and $(x, a) \mapsto D_a(x)$ are definable in the same o-minimal structure.

Main result

- $f_a \to F$ uniformly on compact sets.
- For each $a > 0$, D_a is a conservative field of f_a. (e.g. $D_a = \nabla f_a$ or ∂f_a).
- The maps $(x, a) \mapsto f_a(x)$ and $(x, a) \mapsto D_a(x)$ are definable in the same o-minimal structure.

Theorem

D_F is a conservative set-valued field of F.

$$D_F(x) = \{v \in \mathbb{R}^d : (x_n, v_n, a_n) \to (x, v, 0) \text{ with } v_n \in D_{a_n}(x_n)\}$$

Schechtman, 2024

Main result

- $f_a \to F$ uniformly on compact sets.
- For each $a > 0$, D_a is a conservative field of f_a. (e.g. $D_a = \nabla f_a$ or ∂f_a).
- The maps $(x, a) \mapsto f_a(x)$ and $(x, a) \mapsto D_a(x)$ are definable in the same o-minimal structure.

Theorem

D_F is a conservative set-valued field of F.

$$D_F(x) = \{v \in \mathbb{R}^d : (x_n, v_n, a_n) \to (x, v, 0) \text{ with } v_n \in D_{a_n}(x_n)\}$$

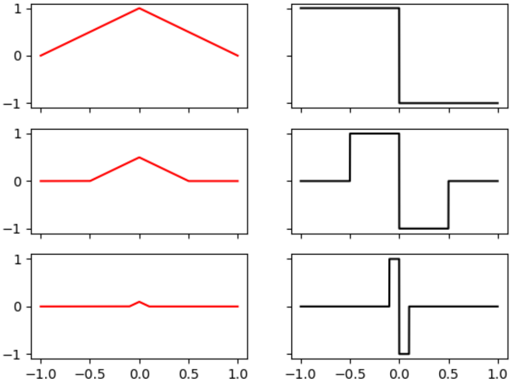
Schechtman, 2024

- D_F can be constructed from an arbitrary family of conservative fields of f_a.
- For almost every $x \in \mathbb{R}^d$, $D_F(x) = \{\nabla F(x)\}$.
- We can partition $\mathbb{R}^d$ into $(\mathcal{M}_i)$ such that for $y \in \mathcal{M}_i$,

$$D_F(y) \subset \nabla_{\mathcal{M}} F(y) + \mathcal{N}_{\mathcal{M}_i}(y)\,.$$

Careful

$D_F \neq \partial F$ even if F and (f_a) are definable and smooth !



$f_a(x)$ and $f_a'(x)$

Careful

It is not sufficient to have f_a definable for each a.

We need the whole map $\quad (x,a) \mapsto f_a(x) \quad$ to be definable!

Careful

It is not sufficient to have f_a definable for each a.

We need the whole map $\quad (x,a) \mapsto f_a(x) \quad$ to be definable!

Recall $f_a(x) = a\sin(x/a)$.

- f_a restricted to $[-1,1]$ is subanalytic.
- The map $(x,a) \mapsto f_a(x)$ is not!

non Lipschitz extensions?

- New proof, allowing non Lipschitz f_a.
  - ➔ Coming soon!
  - ➔ Using the variational stratification as a definition of a conservative field.

Summary

- The subgradients of a definable family of functions converge to a set-valued conservative field.
  - Similar result for conservative Jacobians.
- Smoothing methods converge to a D_F-critical point.
  - ➔ Meaningful first-order optimality condition.
- Future works. Extension to the non Lipschitz case?

Supplementary material

Clarke's subgradient

$$F : \mathbb{R}^d \to \mathbb{R} \text{ locally Lipschitz continuous}\,.$$

Clarke's subgradient

$$F : \mathbb{R}^d \to \mathbb{R} \text{ locally Lipschitz continuous}\,.$$

$$\partial F(x) := \overline{\text{conv}}\{v \in \mathbb{R}^d : x_n \to x, F \text{ differentiable at } x_n, \nabla F(x_n) \to v\}$$

Clarke's subgradient

$$F : \mathbb{R}^d \to \mathbb{R} \text{ locally Lipschitz continuous}\,.$$

$$\partial F(x) := \overline{\text{conv}}\{v \in \mathbb{R}^d : x_n \to x, F \text{ differentiable at } x_n, \nabla F(x_n) \to v\}$$

Properties

- Clarke's subgradient: $\partial F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$, i.e. $\partial F(x) \subset \mathbb{R}^d$.

Clarke's subgradient

$$F : \mathbb{R}^d \to \mathbb{R} \text{ locally Lipschitz continuous}\,.$$

$$\partial F(x) := \overline{\mathrm{conv}}\{v \in \mathbb{R}^d : x_n \to x, F \text{ differentiable at } x_n, \nabla F(x_n) \to v\}$$

Properties

- Clarke's subgradient: $\partial F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$, i.e. $\partial F(x) \subset \mathbb{R}^d$.
- If F is C^1 around x, then $\partial F(x) = \{\nabla F(x)\}$.

Clarke's subgradient

$$F : \mathbb{R}^d \to \mathbb{R} \text{ locally Lipschitz continuous}\,.$$

$$\partial F(x) := \overline{\mathrm{conv}}\{v \in \mathbb{R}^d : x_n \to x, F \text{ differentiable at } x_n, \nabla F(x_n) \to v\}$$

Properties

- Clarke's subgradient: $\partial F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$, i.e. $\partial F(x) \subset \mathbb{R}^d$.
- If F is C^1 around x, then $\partial F(x) = \{\nabla F(x)\}$.
- If F is convex, then ∂F is the convex subgradient.

Clarke's subgradient

$$F : \mathbb{R}^d \to \mathbb{R} \text{ locally Lipschitz continuous}\,.$$

$$\partial F(x) := \overline{\text{conv}}\{v \in \mathbb{R}^d : x_n \to x, F \text{ differentiable at } x_n, \nabla F(x_n) \to v\}$$

Properties

- Clarke's subgradient: $\partial F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$, i.e. $\partial F(x) \subset \mathbb{R}^d$.
- If F is C^1 around x, then $\partial F(x) = \{\nabla F(x)\}$.
- If F is convex, then ∂F is the convex subgradient.
- Optimality condition. If x^* a local minimum, then

$$0 \in \partial F(x^*)\,.$$

Sketch of proof

Goal.

$$F(\mathsf{x}(1)) - F(\mathsf{x}(0)) = \int_0^1 \langle \mathsf{v}(t), \dot{\mathsf{x}}(t) \rangle \, \mathrm{d}t \qquad \text{with } \mathsf{v}(t) \in D_F(\mathsf{x}(t)) \, .$$

Sketch of proof

Goal.

$$F(\mathsf{x}(1)) - F(\mathsf{x}(0)) = \int_0^1 \langle \mathsf{v}(t), \dot{\mathsf{x}}(t) \rangle \, \mathrm{d}t \qquad \text{with } \mathsf{v}(t) \in D_F(\mathsf{x}(t))\,.$$

1) **One dimensional version.** If $\mathsf{x}_a(t) \to \mathsf{x}(t)$, then $\dot{\mathsf{x}}_a(t) \to \dot{\mathsf{x}}(t)$ almost everywhere.

Sketch of proof

Goal.

$$F(\mathsf{x}(1)) - F(\mathsf{x}(0)) = \int_0^1 \langle \mathsf{v}(t), \dot{\mathsf{x}}(t) \rangle \, \mathrm{d}t \qquad \text{with } \mathsf{v}(t) \in D_F(\mathsf{x}(t)) \,.$$

1) **One dimensional version.** If $\mathsf{x}_a(t) \to \mathsf{x}(t)$, then $\dot{\mathsf{x}}_a(t) \to \dot{\mathsf{x}}(t)$ almost everywhere.
2) **Approximating family.** Construct $\mathsf{x}_a(t), \mathsf{v}_a(t) \to \mathsf{x}(t), \mathsf{v}(t)$ such that $\mathsf{v}_a(t) \in D_a(\mathsf{x}_a(t))$ a.e.

Sketch of proof

Goal.

$$F(\mathsf{x}(1)) - F(\mathsf{x}(0)) = \int_0^1 \langle \mathsf{v}(t), \dot{\mathsf{x}}(t)\rangle \,\mathrm{d}t \qquad \text{with } \mathsf{v}(t) \in D_F(\mathsf{x}(t))\,.$$

1) **One dimensional version.** If $\mathsf{x}_a(t) \to \mathsf{x}(t)$, then $\dot{\mathsf{x}}_a(t) \to \dot{\mathsf{x}}(t)$ almost everywhere.
2) **Approximating family.** Construct $\mathsf{x}_a(t), \mathsf{v}_a(t) \to \mathsf{x}(t), \mathsf{v}(t)$ such that $\mathsf{v}_a(t) \in D_a(\mathsf{x}_a(t))$ a.e.
3) **Thus.**

$$f_a(\mathsf{x}_a(1)) - f_a(\mathsf{x}_a(0)) = \int_0^1 \langle \dot{\mathsf{x}}_a(t), \mathsf{v}_a(t)\rangle \,\mathrm{d}t$$

Sketch of proof

Goal.

$$F(\mathsf{x}(1)) - F(\mathsf{x}(0)) = \int_0^1 \langle \mathsf{v}(t), \dot{\mathsf{x}}(t) \rangle \, \mathrm{d}t \qquad \text{with } \mathsf{v}(t) \in D_F(\mathsf{x}(t)) \,.$$

1) **One dimensional version.** If $\mathsf{x}_a(t) \to \mathsf{x}(t)$, then $\dot{\mathsf{x}}_a(t) \to \dot{\mathsf{x}}(t)$ almost everywhere.
2) **Approximating family.** Construct $\mathsf{x}_a(t), \mathsf{v}_a(t) \to \mathsf{x}(t), \mathsf{v}(t)$ such that $\mathsf{v}_a(t) \in D_a(\mathsf{x}_a(t))$ a.e.
3) **Thus.**

$$f_a(\mathsf{x}_a(1)) - f_a(\mathsf{x}_a(0)) = \int_0^1 \langle \dot{\mathsf{x}}_a(t), \mathsf{v}_a(t) \rangle \, \mathrm{d}t$$

passing to the limit

$$F(\mathsf{x}(1)) - F(\mathsf{x}(0)) = \int_0^1 \langle \dot{\mathsf{x}}(t), \mathsf{v}(t) \rangle \, \mathrm{d}t \,.$$