Variational analysis with smooth substructure

Dmitriy Drusvyatskiy,
School of ORIE, Cornell University

January 17, 2012

Goals

- Intuitive notion of identifiable sets.
  - Existence, calculus.
  - Properties of identifiable sets.
- Semi-algebraic geometry.
  - Identifiable manifolds exist generically.
  - Size of semi-algebraic subdifferential graphs.

Preliminaries (Subdifferentials)

For a function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$, a vector $\bar{v}$ is a Frechét subgradient at $\bar{x}$, denoted $\bar{v} \in \hat{\partial} f(\bar{x})$, if

$$f(x) \geq f(\bar{x}) + \langle \bar{v}, x - \bar{x} \rangle + o(|x - \bar{x}|).$$

Preliminaries (Subdifferentials)

For a function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$, a vector $\bar{v}$ is a Frechét subgradient at $\bar{x}$, denoted $\bar{v} \in \hat{\partial} f(\bar{x})$, if

$$f(x) \geq f(\bar{x}) + \langle \bar{v}, x - \bar{x} \rangle + o(|x - \bar{x}|).$$

The limiting subdifferential at $\bar{x}$ is

$$\partial f(\bar{x}) = \{ \lim_{i \to \infty} v_i : v_i \in \hat{\partial} f(x_i), x_i \to \bar{x}, f(x_i) \to f(\bar{x}) \}.$$

Preliminaries (Subdifferentials)

For a function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$, a vector $\bar{v}$ is a Frechét subgradient at $\bar{x}$, denoted $\bar{v} \in \hat{\partial} f(\bar{x})$, if

$$f(x) \geq f(\bar{x}) + \langle \bar{v}, x - \bar{x} \rangle + o(|x - \bar{x}|).$$

The limiting subdifferential at $\bar{x}$ is

$$\partial f(\bar{x}) = \{ \lim_{i \to \infty} v_i : v_i \in \hat{\partial} f(x_i), x_i \to \bar{x}, f(x_i) \to f(\bar{x}) \}.$$

Definition (Generalized critical points)

$\bar{x}$ is a critical point of f if $0 \in \partial f(\bar{x})$.

Preliminaries (Subdifferentials)

For a function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$, a vector $\bar{v}$ is a Frechét subgradient at $\bar{x}$, denoted $\bar{v} \in \hat{\partial} f(\bar{x})$, if

$$f(x) \geq f(\bar{x}) + \langle \bar{v}, x - \bar{x} \rangle + o(|x - \bar{x}|).$$

The limiting subdifferential at $\bar{x}$ is

$$\partial f(\bar{x}) = \{ \lim_{i \to \infty} v_i : v_i \in \hat{\partial} f(x_i), x_i \to \bar{x}, f(x_i) \to f(\bar{x}) \}.$$

Definition (Generalized critical points)

$\bar{x}$ is a critical point of f if $0 \in \partial f(\bar{x})$.

- For convex f, critical points are global minimizers.
- If f is $\mathbf{C}^1$-smooth, criticality reduces to the classical condition $\nabla f(x) = 0$.

Motivation (Sensitivity Analysis)

Consider the perturbed functions

$$f_v(x) = f(x) - \langle v, x \rangle. \qquad [\textit{For simplicity}],$$

and suppose $\bar{x}$ is critical for $f_{\bar{v}}$, that is $\bar{v} \in \partial f(\bar{x})$.

Motivation (Sensitivity Analysis)

Consider the perturbed functions

$$f_v(x) = f(x) - \langle v, x \rangle. \qquad [\textit{For simplicity}],$$

and suppose $\bar{x}$ is critical for $f_{\bar{v}}$, that is $\bar{v} \in \partial f(\bar{x})$.
Sensitivity question: How do critical points of f_v, near $\bar{x}$, behave as v varies near $\bar{v}$?

Motivation (Sensitivity Analysis)

Consider the perturbed functions

$$f_v(x) = f(x) - \langle v, x \rangle. \qquad [\textit{For simplicity}],$$

and suppose $\bar{x}$ is critical for $f_{\bar{v}}$, that is $\bar{v} \in \partial f(\bar{x})$.
Sensitivity question: How do critical points of f_v, near $\bar{x}$, behave as v varies near $\bar{v}$?
Thus given $\bar{v} \in \partial f(\bar{x})$, we want to understand how solutions x_v of

$$v \in \partial f(x),$$

vary, as we perturb v near $\bar{v}$.

Motivating Example

Motivating example

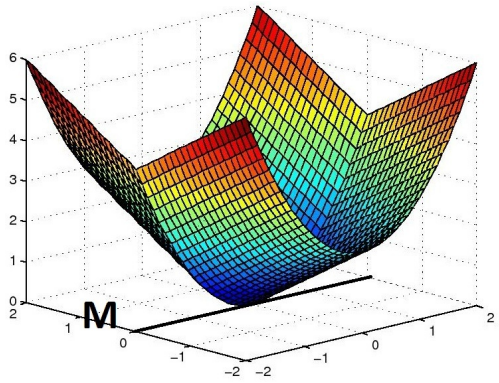


Figure: $f(x, y) = x^2 + |y|, \quad M = \{(t, 0) : -1 < t < 1\}$

Motivating Example

Motivating example

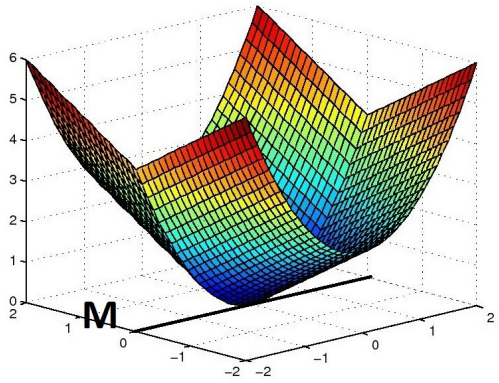


Figure: $f(x, y) = x^2 + |y|, \quad M = \{(t, 0) : -1 < t < 1\}$

- Observe $(0, 0) \in \partial f(0, 0)$.

Motivating Example

Motivating example

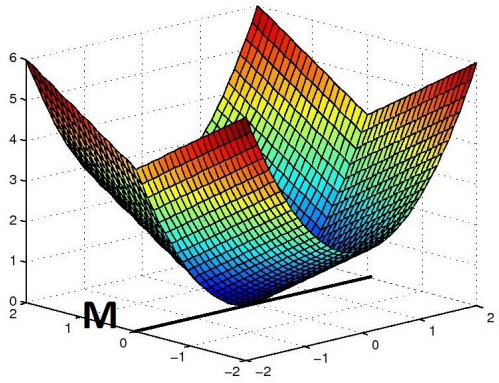


Figure: $f(x,y) = x^2 + |y|, \quad M = \{(t,0) : -1 < t < 1\}$

- Observe $(0,0) \in \partial f(0,0)$.

All perturbed solutions x_v of $v \in \partial f(x)$ lie on $M \implies M$ captures all the sensitivity information!

Motivating Example

Motivating example

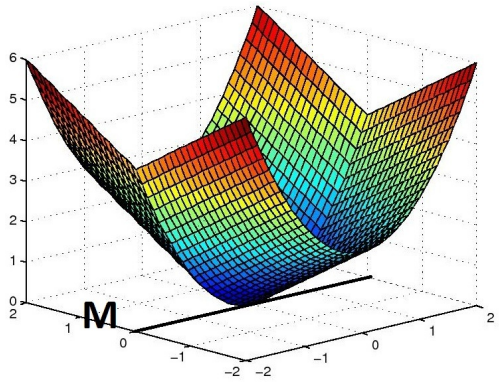


Figure: $f(x,y) = x^2 + |y|, \quad M = \{(t,0) : -1 < t < 1\}$

- Observe $(0,0) \in \partial f(0,0)$.

All perturbed solutions x_v of $v \in \partial f(x)$ lie on $M \implies M$ captures all the sensitivity information!

- Only the restriction $f\big|_M$ matters!

Motivating Example

Motivating example

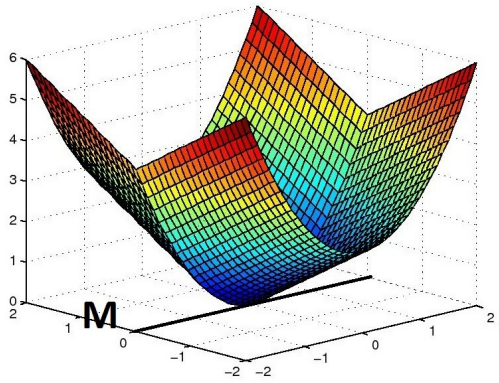


Figure: $f(x,y) = x^2 + |y|, \quad M = \{(t,0) : -1 < t < 1\}$

- Observe $(0,0) \in \partial f(0,0)$.

All perturbed solutions x_v of $v \in \partial f(x)$ lie on $M \implies M$ captures all the sensitivity information!

- Only the restriction $f\big|_M$ matters!
- Goal: Look for small, well-behaved sets capturing only the essential information.

Finite identification

Consider the system

$$v \in \partial f(x).$$

Finite identification

Consider the system

$$v \in \partial f(x).$$

Definition (Identifiable sets)

A set $M \subset \mathbf{R}^n$ is identifiable at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if

$$\left.\begin{array}{c} x_i \to \bar{x}, v_i \to \bar{v} \\ v_i \in \partial f(x_i) \end{array}\right\} \Longrightarrow x_i \in M \text{ for all large } i,$$

Finite identification

Consider the system

$$v \in \partial f(x).$$

Definition (Identifiable sets)

A set $M \subset \mathbf{R}^n$ is identifiable at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if

$$\left.\begin{array}{c} x_i \to \bar{x}, v_i \to \bar{v} \\ v_i \in \partial f(x_i) \end{array}\right\} \Longrightarrow x_i \in M \text{ for all large } i,$$

Example (Normal cone map)

Let $\partial f = N_Q$ for a cube $Q \subset \mathbf{R}^3$.

Finite identification

Consider the system

$$v \in \partial f(x).$$

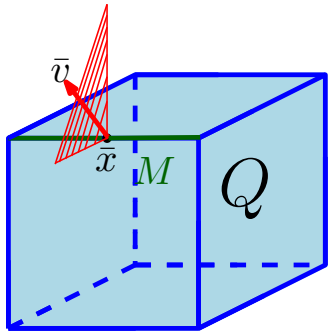
Definition (Identifiable sets)

A set $M \subset \mathbf{R}^n$ is identifiable at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if

$$\left.\begin{array}{r} x_i \to \bar{x}, v_i \to \bar{v} \\ v_i \in \partial f(x_i) \end{array}\right\} \Longrightarrow x_i \in M \text{ for all large } i,$$

Example (Normal cone map)

Let $\partial f = N_Q$ for a cube $Q \subset \mathbf{R}^3$.



Finite identification

Consider the system

$$v \in \partial f(x).$$

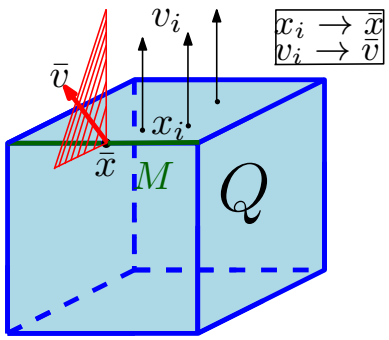
Definition (Identifiable sets)

A set $M \subset \mathbf{R}^n$ is identifiable at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if

$$\left.\begin{array}{r} x_i \to \bar{x}, v_i \to \bar{v} \\ v_i \in \partial f(x_i) \end{array}\right\} \Longrightarrow x_i \in M \text{ for all large } i,$$

Example (Normal cone map)

Let $\partial f = N_Q$ for a cube $Q \subset \mathbf{R}^3$.



Finite identification

Consider the system

$$v \in \partial f(x).$$

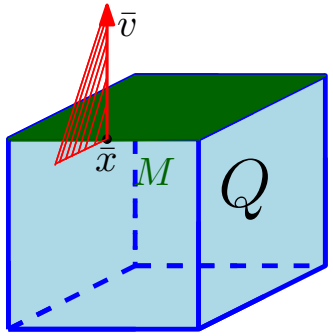
Definition (Identifiable sets)

A set $M \subset \mathbf{R}^n$ is identifiable at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if

$$\left.\begin{array}{r} x_i \to \bar{x}, v_i \to \bar{v} \\ v_i \in \partial f(x_i) \end{array}\right\} \Longrightarrow x_i \in M \text{ for all large } i,$$

Example (Normal cone map)

Let $\partial f = N_Q$ for a cube $Q \subset \mathbf{R}^3$.



Finite identification

Consider the system

$$v \in \partial f(x).$$

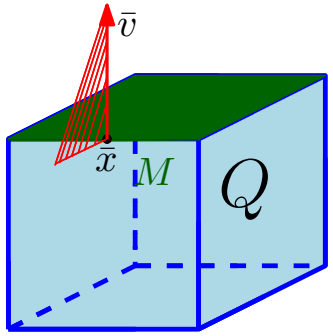
Definition (Identifiable sets)

A set $M \subset \mathbf{R}^n$ is identifiable at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if

$$\left.\begin{array}{r} x_i \to \bar{x}, v_i \to \bar{v} \\ v_i \in \partial f(x_i) \end{array}\right\} \Longrightarrow x_i \in M \text{ for all large } i,$$

Example (Normal cone map)

Let $\partial f = N_Q$ for a cube $Q \subset \mathbf{R}^3$.



In this case $M = \operatorname{argmax}_{x \in Q} \langle \bar{v}, x \rangle$.

Order of growth

Why are identifiable sets interesting?

Order of growth

Why are identifiable sets interesting?

Proposition (D, Lewis)

Suppose M is an identifiable set at $\bar{x}$ for $0 \in \hat{\partial} f(\bar{x})$.

Order of growth

Why are identifiable sets interesting?

Proposition (D, Lewis)

Suppose M is an identifiable set at $\bar{x}$ for $0 \in \hat{\partial} f(\bar{x})$.

- *$\bar{x}$ is a (strict) local minimizer of $f \iff \bar{x}$ is a (strict) local minimizer of f on M.*
- *f grows quadratically near $\bar{x} \iff f$ grows quadratically on M near $\bar{x}$.*

Subgradient dynamical systems

Consider a subgradient dynamical system

$$\dot{x}(t) \in \partial f(x(t)) \quad \text{(a.e.)}$$

where $f : \mathbf{R}^n \to \overline{\mathbf{R}}$.

Subgradient dynamical systems

Consider a subgradient dynamical system

$$\dot{x}(t) \in \partial f(x(t)) \quad \text{(a.e.)}$$

where $f : \mathbf{R}^n \to \overline{\mathbf{R}}$.
Suppose

Subgradient dynamical systems

Consider a subgradient dynamical system

$$\dot{x}(t) \in \partial f(x(t)) \quad \text{(a.e.)}$$

where $f : \mathbf{R}^n \to \overline{\mathbf{R}}$.
Suppose

- $x : [0, \infty) \to \mathbf{R}^n$ is a trajectory of finite length converging to a critical point $\bar{x}$.
- M is identifiable at $\bar{x}$ for 0.

Subgradient dynamical systems

Consider a subgradient dynamical system

$$\dot{x}(t) \in \partial f(x(t)) \quad \text{(a.e.)}$$

where $f : \mathbf{R}^n \to \overline{\mathbf{R}}$.
Suppose

- $x : [0, \infty) \to \mathbf{R}^n$ is a trajectory of finite length converging to a critical point $\bar{x}$.
- M is identifiable at $\bar{x}$ for 0.

Then $x(t) \in M$ for a.e. $t > T$, for some $T > 0$.

Subgradient dynamical systems

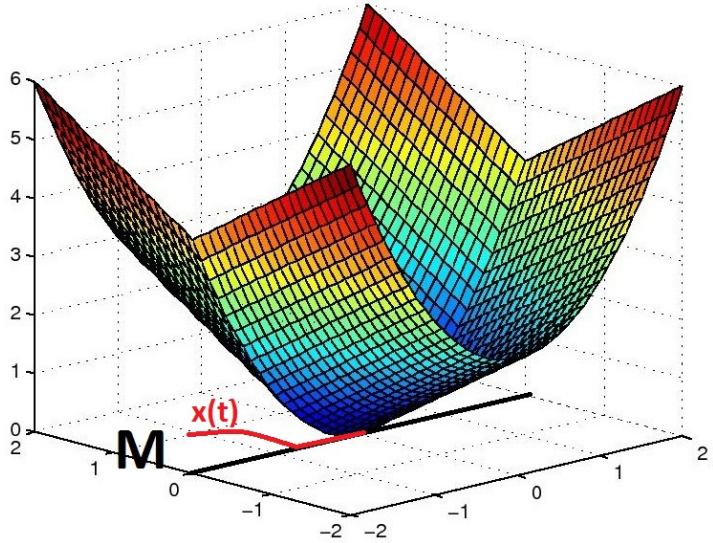
Consider a subgradient dynamical system

$$\dot{x}(t) \in \partial f(x(t)) \quad \text{(a.e.)}$$

where $f : \mathbf{R}^n \to \overline{\mathbf{R}}$.
Suppose

- $x : [0, \infty) \to \mathbf{R}^n$ is a trajectory of finite length converging to a critical point $\bar{x}$.
- M is identifiable at $\bar{x}$ for 0.

Then $x(t) \in M$ for a.e. $t > T$, for some $T > 0$.



Locally minimal identifiable sets

Clearly all of $\mathbf{R}^n$ is identifiable at $\bar{x}$ for $\bar{v}$ (not interesting). So...

Locally minimal identifiable sets

Clearly all of $\mathbf{R}^n$ is identifiable at $\bar{x}$ for $\bar{v}$ (not interesting). So...
Question: What are the smallest possible identifiable sets?

Locally minimal identifiable sets

Clearly all of $\mathbf{R}^n$ is identifiable at $\bar{x}$ for $\bar{v}$ (not interesting). So...
Question: What are the smallest possible identifiable sets?

Definition

An identifiable set M at $\bar{x}$ for $\bar{v}$ is locally minimal if

$$M' \text{ identifiable at } \bar{x} \text{ for } \bar{v} \Longrightarrow M \subset M', \text{ locally near } \bar{x}.$$

Existence and calculus

Locally minimal identifiable sets exist for

Existence and calculus

Locally minimal identifiable sets exist for

- convex polyhedra $[M = \operatorname{argmax}_{x \in Q} \langle \bar{v}, x \rangle]$,
- piecewise quadratic functions,
- max-type functions: $f(x) = \max\{g_1(x), \ldots, g_k(x)\}$ for $\mathbf{C}^1$-smooth g_i,
- fully amenable functions: $f(x) = g(F(x))$ where
  1. F is $\mathbf{C}^2$-smooth,
  2. g is (convex) piecewise quadratic,
  3. qualification condition holds.

Existence and calculus

Locally minimal identifiable sets exist for

- convex polyhedra $[M = \operatorname{argmax}_{x \in Q} \langle \bar{v}, x \rangle]$,
- piecewise quadratic functions,
- max-type functions: $f(x) = \max\{g_1(x), \ldots, g_k(x)\}$ for $\mathbf{C}^1$-smooth g_i,
- fully amenable functions: $f(x) = g(F(x))$ where
  1. F is $\mathbf{C}^2$-smooth,
  2. g is (convex) piecewise quadratic,
  3. qualification condition holds.

A strong chain rule is available for composite functions

$$f(x) = g(F(x)).$$

Topological regularity

Intriguing Question: Do locally minimal identifiable sets exist for any convex function?

Topological regularity

Intriguing Question: Do locally minimal identifiable sets exist for any convex function? No.

Topological regularity

Intriguing Question: Do locally minimal identifiable sets exist for any convex function? No.

This will follow from the following observation.

Topological regularity

Intriguing Question: Do locally minimal identifiable sets exist for any convex function? No.
This will follow from the following observation.

Proposition (Topological regularity)

Consider decreasing sequence of open neighbourhoods

$$V_1 \supset V_2 \supset V_3 \supset \ldots, \text{ with } V_i \downarrow \{\bar{v}\},$$

and

Topological regularity

Intriguing Question: Do locally minimal identifiable sets exist for any convex function? No.
This will follow from the following observation.

Proposition (Topological regularity)

Consider decreasing sequence of open neighbourhoods

$$V_1 \supset V_2 \supset V_3 \supset \ldots, \text{ with } V_i \downarrow \{\bar{v}\},$$

and

$$(\partial f)^{-1}(V_1) \supset (\partial f)^{-1}(V_2) \supset (\partial f)^{-1}(V_3) \supset \ldots.$$

Topological regularity

Intriguing Question: Do locally minimal identifiable sets exist for any convex function? No.
This will follow from the following observation.

Proposition (Topological regularity)

Consider decreasing sequence of open neighbourhoods

$$V_1 \supset V_2 \supset V_3 \supset \ldots, \text{ with } V_i \downarrow \{\bar{v}\},$$

and

$$(\partial f)^{-1}(V_1) \supset (\partial f)^{-1}(V_2) \supset (\partial f)^{-1}(V_3) \supset \ldots.$$

Then M is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$ $\Longleftrightarrow$

$$M = (\partial f)^{-1}(V_i) = (\partial f)^{-1}(V_j) \textit{ locally near } \bar{x}$$

for all large i, j.

Topological regularity

Intriguing Question: Do locally minimal identifiable sets exist for any convex function? No.
This will follow from the following observation.

Proposition (Topological regularity)

Consider decreasing sequence of open neighbourhoods

$$V_1 \supset V_2 \supset V_3 \supset \ldots, \text{ with } V_i \downarrow \{\bar{v}\},$$

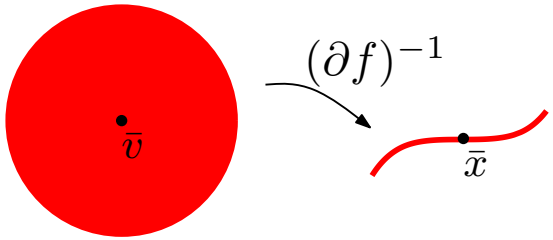
and

$$(\partial f)^{-1}(V_1) \supset (\partial f)^{-1}(V_2) \supset (\partial f)^{-1}(V_3) \supset \ldots.$$

Then M is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$ $\iff$

$$M = (\partial f)^{-1}(V_i) = (\partial f)^{-1}(V_j) \textit{ locally near } \bar{x}$$

for all large i, j.



Topological regularity

Intriguing Question: Do locally minimal identifiable sets exist for any convex function? No.
This will follow from the following observation.

Proposition (Topological regularity)

Consider decreasing sequence of open neighbourhoods

$$V_1 \supset V_2 \supset V_3 \supset \ldots, \text{ with } V_i \downarrow \{\bar{v}\},$$

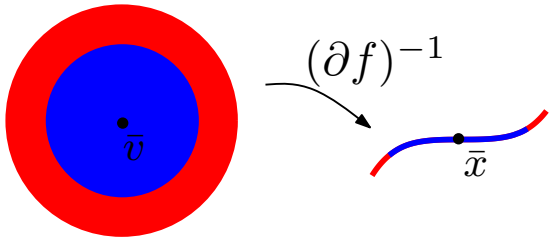
and

$$(\partial f)^{-1}(V_1) \supset (\partial f)^{-1}(V_2) \supset (\partial f)^{-1}(V_3) \supset \ldots.$$

Then M is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$ $\iff$

$$M = (\partial f)^{-1}(V_i) = (\partial f)^{-1}(V_j) \textit{ locally near } \bar{x}$$

for all large i, j.



Topological regularity

Intriguing Question: Do locally minimal identifiable sets exist for any convex function? No.
This will follow from the following observation.

Proposition (Topological regularity)

Consider decreasing sequence of open neighbourhoods

$$V_1 \supset V_2 \supset V_3 \supset \ldots, \text{ with } V_i \downarrow \{\bar{v}\},$$

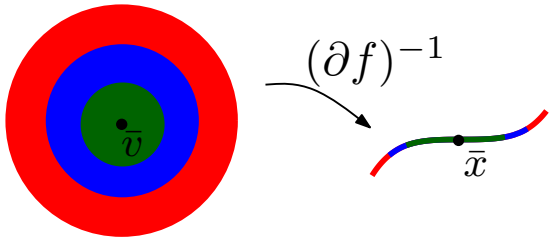
and

$$(\partial f)^{-1}(V_1) \supset (\partial f)^{-1}(V_2) \supset (\partial f)^{-1}(V_3) \supset \ldots.$$

Then M is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$ $\iff$

$$M = (\partial f)^{-1}(V_i) = (\partial f)^{-1}(V_j) \textit{ locally near } \bar{x}$$

for all large i, j.



Existence and calculus

Convex functions may fail to admit locally minimal identifiable sets!

Existence and calculus

Convex functions may fail to admit locally minimal identifiable sets!

Example

Existence and calculus

Convex functions may fail to admit locally minimal identifiable sets!

Example

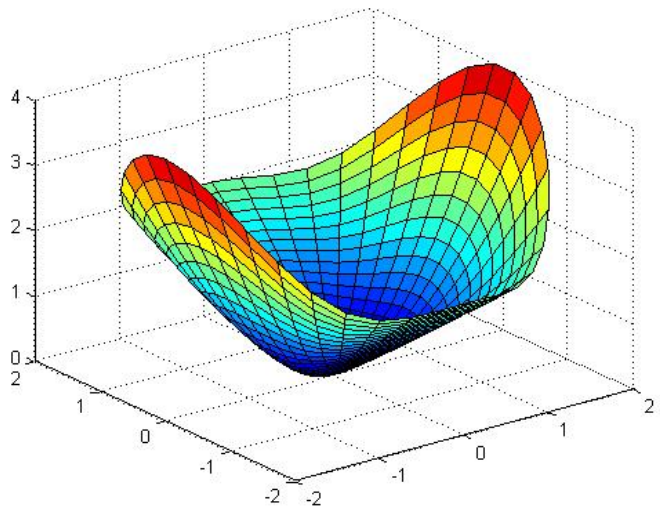


Figure: $f(x, y) = \sqrt{x^4 + y^2}$.

Existence and calculus

Convex functions may fail to admit locally minimal identifiable sets!

Example

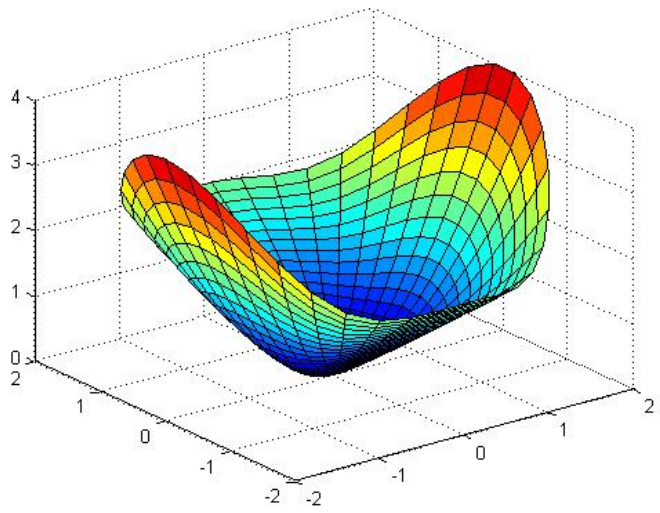


Figure: $f(x, y) = \sqrt{x^4 + y^2}$.

Existence and calculus

Convex functions may fail to admit locally minimal identifiable sets!

Example

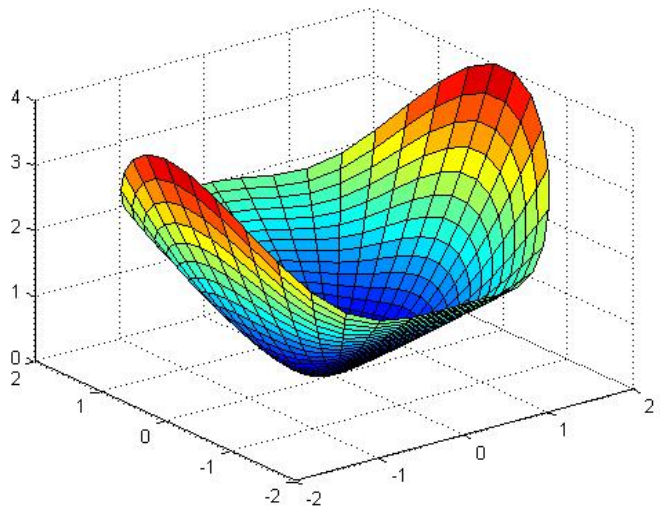


Figure: $f(x, y) = \sqrt{x^4 + y^2}$.

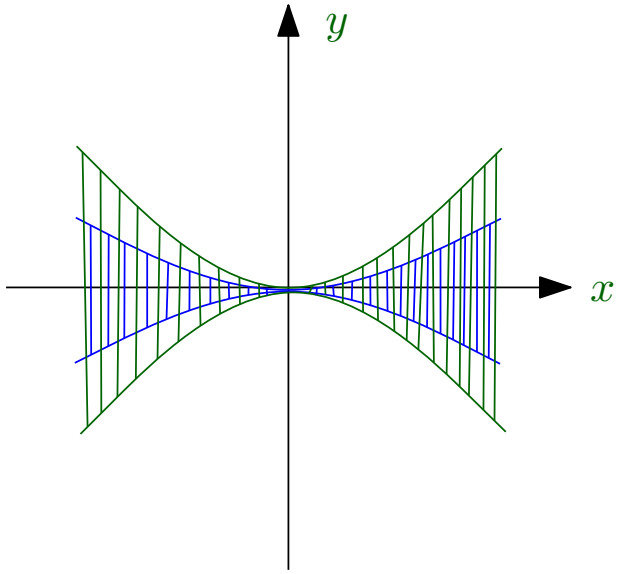


Figure: Level sets: $[|\nabla f| < \epsilon]$.

Existence and calculus

Convex functions may fail to admit locally minimal identifiable sets!

Example

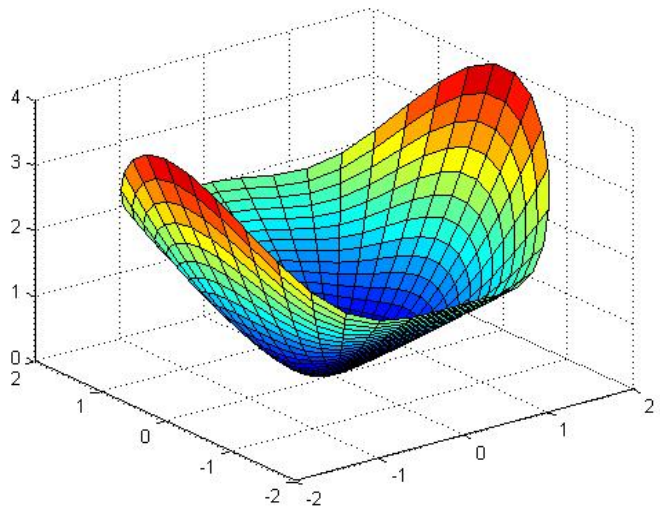


Figure: $f(x, y) = \sqrt{x^4 + y^2}$.

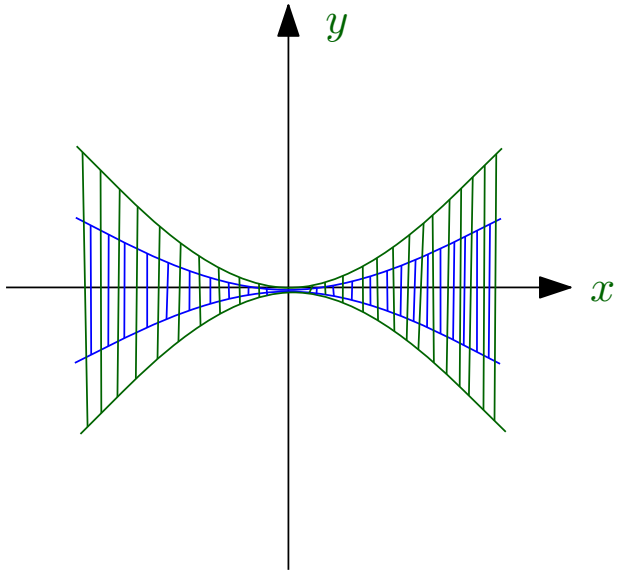


Figure: Level sets: $[|\nabla f| < \epsilon]$.

Level sets of $|\nabla f|$ get "pinched".

Existence and calculus

Convex functions may fail to admit locally minimal identifiable sets!

Example

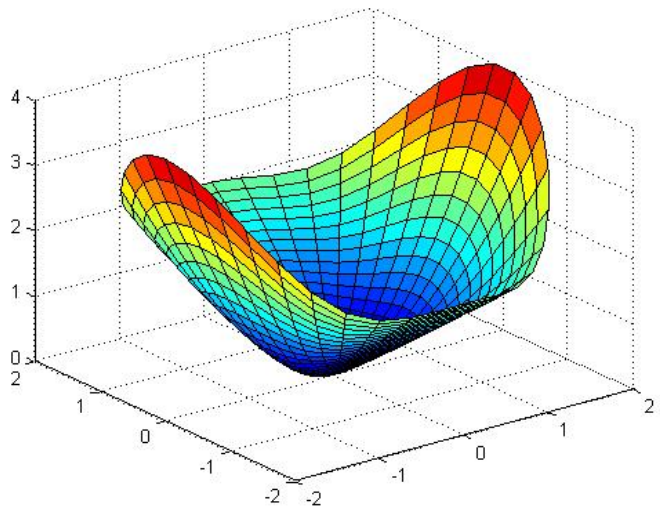


Figure: $f(x,y) = \sqrt{x^4 + y^2}$.

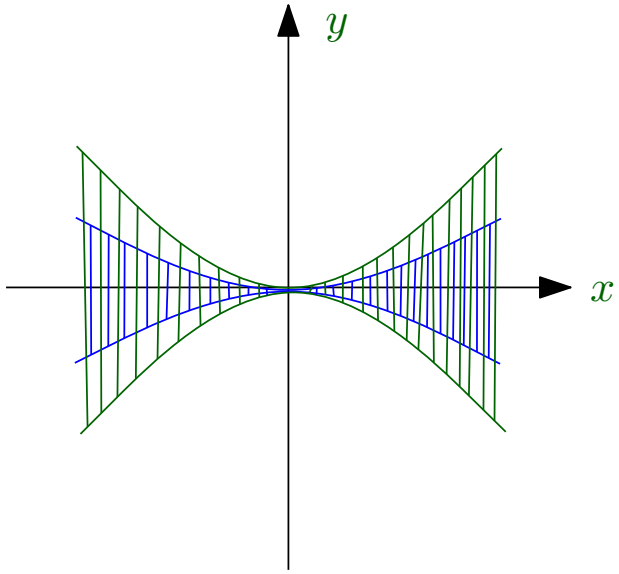


Figure: Level sets: $[|\nabla f| < \epsilon]$.

Level sets of $|\nabla f|$ get "pinched". No locally minimal identifiable set at $\bar{x} = (0,0)$ for $\bar{v} = (0,0)$.

Relation to critical cones

$Q \subset \mathbf{R}^n$ that is Clarke regular at $\bar{x}$ if

$$\hat{N}_Q(\bar{x}) = N_Q(\bar{x}).$$

Relation to critical cones

$Q \subset \mathbf{R}^n$ that is Clarke regular at $\bar{x}$ if

$$\hat{N}_Q(\bar{x}) = N_Q(\bar{x}).$$

The tangent cone is

$$T_Q(\bar{x}) := \{\lim_{i\to\infty} \frac{x_i - \bar{x}}{\tau_i} : x_i \xrightarrow{Q} \bar{x}, \tau_i \to 0\}.$$

Relation to critical cones

$Q \subset \mathbf{R}^n$ that is Clarke regular at $\bar{x}$ if

$$\hat{N}_Q(\bar{x}) = N_Q(\bar{x}).$$

The tangent cone is

$$T_Q(\bar{x}) := \{\lim_{i\to\infty} \frac{x_i - \bar{x}}{\tau_i} : x_i \xrightarrow{Q} \bar{x}, \tau_i \to 0\}.$$

The critical cone of Q at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$ is

$$K_Q(\bar{x}, \bar{v}) := T_Q(\bar{x}) \cap \bar{v}^{\perp}.$$

Relation to critical cones

$Q \subset \mathbf{R}^n$ that is Clarke regular at $\bar{x}$ if

$$\hat{N}_Q(\bar{x}) = N_Q(\bar{x}).$$

The tangent cone is

$$T_Q(\bar{x}) := \{\lim_{i\to\infty} \frac{x_i - \bar{x}}{\tau_i} : x_i \xrightarrow{Q} \bar{x}, \tau_i \to 0\}.$$

The critical cone of Q at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$ is

$$K_Q(\bar{x}, \bar{v}) := T_Q(\bar{x}) \cap \bar{v}^{\perp}.$$

Critical cones are crucial for analysing polyhedral variational inequalities

$$0 \in F(x, p) + N_{S(p)}(x),$$

where $S(p)$ are convex polyhedra,

Relation to critical cones

$Q \subset \mathbf{R}^n$ that is Clarke regular at $\bar{x}$ if

$$\hat{N}_Q(\bar{x}) = N_Q(\bar{x}).$$

The tangent cone is

$$T_Q(\bar{x}) := \{\lim_{i\to\infty} \frac{x_i - \bar{x}}{\tau_i} : x_i \xrightarrow{Q} \bar{x}, \tau_i \to 0\}.$$

The critical cone of Q at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$ is

$$K_Q(\bar{x}, \bar{v}) := T_Q(\bar{x}) \cap \bar{v}^{\perp}.$$

Critical cones are crucial for analysing polyhedral variational inequalities

$$0 \in F(x, p) + N_{S(p)}(x),$$

where $S(p)$ are convex polyhedra, because of

Relation to critical cones

Proposition (Key Property)

If Q is polyhedral, then

$$\operatorname{gph} N_Q = \operatorname{gph} N_{\bar{x}+K_Q(\bar{x},\bar{v})} \text{ locally near } (\bar{x}, \bar{v}).$$

Relation to critical cones

Proposition (Key Property)

If Q is polyhedral, then

$$\operatorname{gph} N_Q = \operatorname{gph} N_{\bar{x}+K_Q(\bar{x},\bar{v})} \text{ locally near } (\bar{x}, \bar{v}).$$

Not true at all beyond polyhedral sets, but

Relation to critical cones

Proposition (Key Property)

If Q is polyhedral, then

$$\operatorname{gph} N_Q = \operatorname{gph} N_{\bar{x}+K_Q(\bar{x},\bar{v})} \text{ locally near } (\bar{x},\bar{v}).$$

Not true at all beyond polyhedral sets, but

Proposition (D, Lewis)

Let M be a (prox-regular) locally minimal identifiable set at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$. Then

$$\operatorname{gph} N_Q = \operatorname{gph} N_M \text{ locally near } (\bar{x},\bar{v}),$$

Relation to critical cones

Proposition (Key Property)

If Q is polyhedral, then

$$\operatorname{gph} N_Q = \operatorname{gph} N_{\bar{x}+K_Q(\bar{x},\bar{v})} \text{ locally near } (\bar{x},\bar{v}).$$

Not true at all beyond polyhedral sets, but

Proposition (D, Lewis)

Let M be a (prox-regular) locally minimal identifiable set at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$. Then

$$\operatorname{gph} N_Q = \operatorname{gph} N_M \text{ locally near } (\bar{x},\bar{v}),$$

and furthermore

$$K(\bar{x},\bar{v}) = \operatorname{cl}\operatorname{conv} T_M(\bar{x}).$$

Relation to critical cones

Proposition (Key Property)

If Q is polyhedral, then

$$\operatorname{gph} N_Q = \operatorname{gph} N_{\bar{x}+K_Q(\bar{x},\bar{v})} \text{ locally near } (\bar{x},\bar{v}).$$

Not true at all beyond polyhedral sets, but

Proposition (D, Lewis)

Let M be a (prox-regular) locally minimal identifiable set at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$. Then

$$\operatorname{gph} N_Q = \operatorname{gph} N_M \text{ locally near } (\bar{x},\bar{v}),$$

and furthermore

$$K(\bar{x},\bar{v}) = \operatorname{cl}\operatorname{conv} T_M(\bar{x}).$$

May use this to study nonpolyhedral variational inequalities!

Algorithmic origins

- Intuitive idea of finite identification, in this setting, is old.

Algorithmic origins

- Intuitive idea of finite identification, in this setting, is old.
- Some algorithms for solving $\min_{x \in Q} f(x)$ would stop in finite time (proximal point algorithm Rockafellar '76).

Algorithmic origins

- Intuitive idea of finite identification, in this setting, is old.
- Some algorithms for solving $\min_{x \in Q} f(x)$ would stop in finite time (proximal point algorithm Rockafellar '76).
- Many algorithms would generate iterates that eventually lie on a distinguished subset of Q

Algorithmic origins

- Intuitive idea of finite identification, in this setting, is old.
- Some algorithms for solving $\min_{x \in Q} f(x)$ would stop in finite time (proximal point algorithm Rockafellar '76).
- Many algorithms would generate iterates that eventually lie on a distinguished subset of Q (subgradient projection Calamai-Moré '87, Newton-like methods Burke-Moré '88, stochastic gradient methods Wright '11).

Algorithmic origins

- Intuitive idea of finite identification, in this setting, is old.
- Some algorithms for solving $\min_{x \in Q} f(x)$ would stop in finite time (proximal point algorithm Rockafellar '76).
- Many algorithms would generate iterates that eventually lie on a distinguished subset of Q (subgradient projection Calamai-Moré '87, Newton-like methods Burke-Moré '88, stochastic gradient methods Wright '11).
- One may try to exploit a nice identifiable set, if one exists; perhaps a $\mathbf{C}^2$-manifold.

Algorithmic origins

- Intuitive idea of finite identification, in this setting, is old.
- Some algorithms for solving $\min_{x \in Q} f(x)$ would stop in finite time (proximal point algorithm Rockafellar '76).
- Many algorithms would generate iterates that eventually lie on a distinguished subset of Q (subgradient projection Calamai-Moré '87, Newton-like methods Burke-Moré '88, stochastic gradient methods Wright '11).
- One may try to exploit a nice identifiable set, if one exists; perhaps a $\mathbf{C}^2$-manifold.
- Use first order method and "track" M. Then make a guess and use Newton's method to accelerate (Survey by Sagastizábal '11 in SIAG/OPT Views and News).

Identifiable manifolds

Identifiable manifolds

When there exists an identifiable manifold M (M is identifiable, M is a manifold, and $f\big|_M$ is smooth), things simplify drastically.

Identifiable manifolds

When there exists an identifiable manifold M (M is identifiable, M is a manifold, and $f\big|_M$ is smooth), things simplify drastically.

Proposition (D-Lewis)

Identifiable manifolds $M \subset \operatorname{dom} f$ are automatically locally minimal.

Identifiable manifolds

When there exists an identifiable manifold M (M is identifiable, M is a manifold, and $f\big|_M$ is smooth), things simplify drastically.

Proposition (D-Lewis)

Identifiable manifolds $M \subset \operatorname{dom} f$ are automatically locally minimal.

- Identifiable manifolds provide a refinement of partly smooth manifolds introduced in Lewis '03.

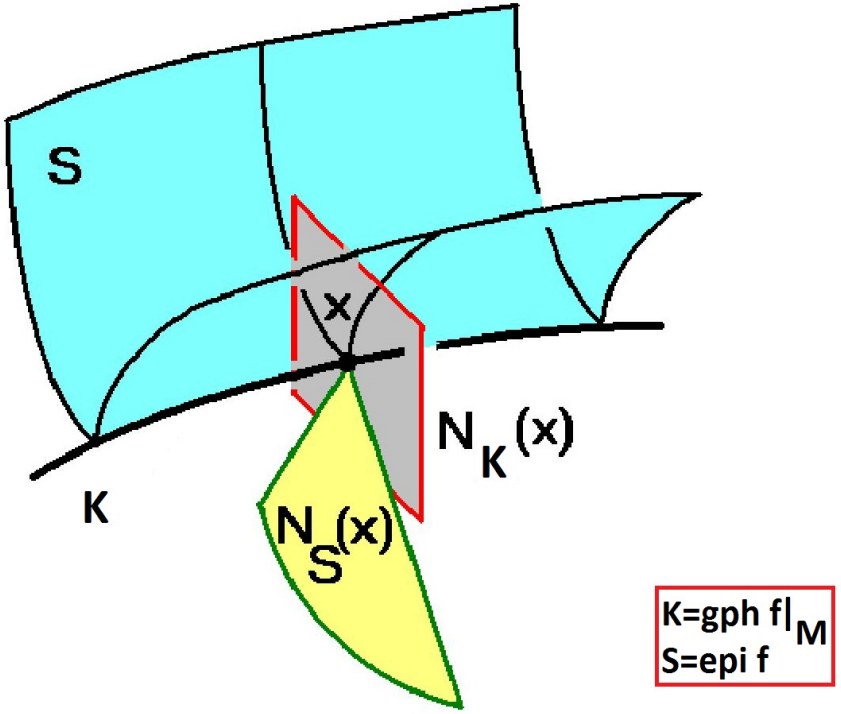
Identifiable manifolds

When there exists an identifiable manifold M (M is identifiable, M is a manifold, and $f\big|_M$ is smooth), things simplify drastically.

Proposition (D-Lewis)

Identifiable manifolds $M \subset \operatorname{dom} f$ are automatically locally minimal.

- Identifiable manifolds provide a refinement of partly smooth manifolds introduced in Lewis '03.



Partly smooth manifolds

Identifiable manifolds axiomatize active manifolds in classical math programming (independent of representation).

Partly smooth manifolds

Identifiable manifolds axiomatize active manifolds in classical math programming (independent of representation).

- When an identifiable manifold exists, nonsmoothness is not intrinsic.

Partly smooth manifolds

Identifiable manifolds axiomatize active manifolds in classical math programming (independent of representation).

- When an identifiable manifold exists, nonsmoothness is not intrinsic.
- So can reduce to the classical setting.

Generic Properties

History: Rockafellar-Spingarn '79, considered problems

$$P(v,u): \quad \min \quad f(x) - \langle v, x \rangle, \\ \text{s.t.} \quad g_i(x) \leq u_i, \text{ for all } i \in I := \{1, \ldots, m\},$$

for smooth f, g_i.

Generic Properties

History: Rockafellar-Spingarn '79, considered problems

$$P(v,u): \quad \min \quad f(x) - \langle v, x \rangle, \\ \text{s.t.} \quad g_i(x) \leq u_i, \text{ for all } i \in I := \{1, \ldots, m\},$$

for smooth f, g_i.

Theorem (Rockafellar-Spingarn '79)

- *For almost all v, active gradients are independent (there exists an active manifold).*
- *For almost all (v,u), the second order conditions hold at every minimizer:*

  Strict complementarity: multipliers are strictly positive and

  Quadratic growth: $f(\cdot) - \langle \bar{v}, \cdot \rangle$ grows quadratically on M near $\bar{x}$.

Generic Properties

How typical are identifiable manifolds? Second-order growth at minimizers?

Generic Properties

How typical are identifiable manifolds? Second-order growth at minimizers?

We answer this question in the setting of semi-algebraic sets:

.

Generic Properties

How typical are identifiable manifolds? Second-order growth at minimizers?

We answer this question in the setting of semi-algebraic sets:

represented as finite union of sets, each defined by finitely many polynomial inequalities.

Generic Properties

How typical are identifiable manifolds? Second-order growth at minimizers?

We answer this question in the setting of semi-algebraic sets:

represented as finite union of sets, each defined by finitely many polynomial inequalities.

Large class of sets for which the word typical has a canonical meaning.

Generic Properties

Theorem (D, Lewis)

Suppose $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ *is semi-algebraic. Consider the perturbed functions*

$$f_p(x) := f(x) + \theta(p, x),$$

and

$$D_{p,x}(x, p) \text{ is } \textit{surjective} \text{ for all } (x, p).$$

Generic Properties

Theorem (D, Lewis)

Suppose $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ *is semi-algebraic. Consider the perturbed functions*

$$f_p(x) := f(x) + \theta(p, x),$$

and

$$D_{p,x}(x, p) \text{ is } \textit{surjective} \text{ for all } (x, p).$$

(e.g. $f_p(x) = f(x) - \langle p, x \rangle$ *or* $f_p(x) = f(x) + \frac{1}{2}|x - p|^2$*).*

Generic Properties

Theorem (D, Lewis)

Suppose $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ is semi-algebraic. Consider the perturbed functions

$$f_p(x) := f(x) + \theta(p, x),$$

and

$$D_{p,x}(x, p) \text{ is } \textit{surjective} \text{ for all } (x, p).$$

(e.g. $f_p(x) = f(x) - \langle p, x \rangle$ or $f_p(x) = f(x) + \frac{1}{2}|x - p|^2$).
Then for a "typical" $p \in P$,

- *f_p has finitely many critical points x_p.*
- *f_p admits an identifiable manifold near each x_p for 0.*
- *every local minimizer x_p of f_p on the manifold is a strong local minimizer of f_p.*
- *Smooth dependence of critical points near x_p on the manifold.*

Dimension

- **Key property of semi-algebraic functions**: Size of subdifferential graphs.

Dimension

- **Key property of semi-algebraic functions**: Size of subdifferential graphs.

Definition

For any semi-algebraic set $A \subset \mathbf{R}^n$, we can write $A = \bigcup_{i=0}^{k} M_i$, where M_i are disjoint semi-algebraic manifolds.

Dimension

- **Key property of semi-algebraic functions**: Size of subdifferential graphs.

Definition

For any semi-algebraic set $A \subset \mathbf{R}^n$, we can write $A = \bigcup_{i=0}^{k} M_i$, where M_i are disjoint semi-algebraic manifolds.

$$\dim A := \max_i \dim M_i.$$

Dimension

- **Key property of semi-algebraic functions**: Size of subdifferential graphs.

Definition

For any semi-algebraic set $A \subset \mathbf{R}^n$, we can write $A = \bigcup_{i=0}^{k} M_i$, where M_i are disjoint semi-algebraic manifolds.

$$\dim A := \max_i \dim M_i.$$

- Dimension is a global property.

Local Dimension

Definition

The local dimension of Q at $\bar{x}$ is

$$\dim_Q(\bar{x}) := \inf_{r>0} \dim(Q \cap B_r(\bar{x})).$$

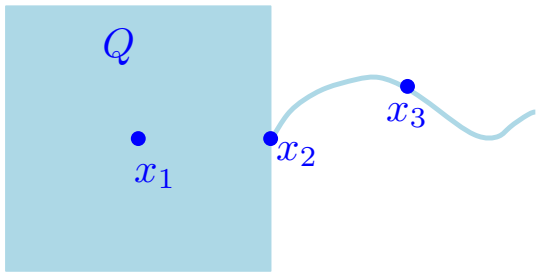
Local Dimension

Definition

The local dimension of Q at $\bar{x}$ is

$$\dim_Q(\bar{x}) := \inf_{r>0} \dim(Q \cap B_r(\bar{x})).$$

Example



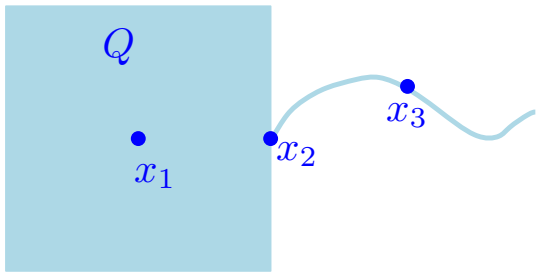
Local Dimension

Definition

The local dimension of Q at $\bar{x}$ is

$$\dim_Q(\bar{x}) := \inf_{r>0} \dim(Q \cap B_r(\bar{x})).$$

Example



$\dim_Q(x_1) = \dim_Q(x_2) = 2$, $\dim_Q(x_3) = 1$

Size of semi-algebraic subdifferential graphs

If $f : \mathbf{R}^n \to \mathbf{R}$ is $\mathbf{C}^1$ smooth, then $\operatorname{gph} \nabla f(x)$ has everywhere n-dimensional graph.

Size of semi-algebraic subdifferential graphs

If $f : \mathbf{R}^n \to \mathbf{R}$ is $\mathbf{C}^1$ smooth, then $\operatorname{gph} \nabla f(x)$ has everywhere n-dimensional graph.

Theorem (Minty '62)

If $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ is convex, ∂f has everywhere n-dimensional graph.
(. . . with computational implications for equations on the graph.)

Size of semi-algebraic subdifferential graphs

If $f : \mathbf{R}^n \to \mathbf{R}$ is $\mathbf{C}^1$ smooth, then $\operatorname{gph} \nabla f(x)$ has everywhere n-dimensional graph.

Theorem (Minty '62)

If $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ is convex, ∂f has everywhere n-dimensional graph.
(. . . with computational implications for equations on the graph.)

- For Lipschitz functions, ∂f usually has a large graph: $2n$-dimensional (Borwein-Wang '00).

Size of semi-algebraic subdifferential graphs

If $f : \mathbf{R}^n \to \mathbf{R}$ is $\mathbf{C}^1$ smooth, then $\operatorname{gph} \nabla f(x)$ has everywhere n-dimensional graph.

Theorem (Minty '62)

If $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ is convex, ∂f has everywhere n-dimensional graph.
(. . . with computational implications for equations on the graph.)

- For Lipschitz functions, ∂f usually has a large graph: $2n$-dimensional (Borwein-Wang '00).

But...

Size of semi-algebraic subdifferential graphs

If $f : \mathbf{R}^n \to \mathbf{R}$ is $\mathbf{C}^1$ smooth, then $\operatorname{gph} \nabla f(x)$ has everywhere n-dimensional graph.

Theorem (Minty '62)

If $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ is convex, ∂f has everywhere n-dimensional graph.
(. . . with computational implications for equations on the graph.)

- For Lipschitz functions, ∂f usually has a large graph: $2n$-dimensional (Borwein-Wang '00).

But...

Theorem (D-Ioffe-Lewis)

If $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ is semi-algebraic, then ∂f has everywhere n-dimensional graph.

Failure for the Clarke Subdifferential

- The theorem is false for the Clarke case!

Failure for the Clarke Subdifferential

- The theorem is false for the Clarke case!
- For Lipschitz f,

$$\partial_c f(\bar{x}) = \mathrm{conv}\,\{\lim_{i\to\infty} \nabla f(x_i) : x_i \to \bar{x}\}.$$

Failure for the Clarke Subdifferential

- The theorem is false for the Clarke case!
- For Lipschitz f,

$$\partial_c f(\bar{x}) = \mathrm{conv}\,\{\lim_{i\to\infty} \nabla f(x_i) : x_i \to \bar{x}\}.$$

- This is counterintuitive since $\hat{\partial} f(\bar{x}) \subset \partial f(\bar{x}) \subset \partial_c f(\bar{x})$.

Failure for the Clarke Subdifferential

- The theorem is false for the Clarke case!
- For Lipschitz f,

$$\partial_c f(\bar{x}) = \mathrm{conv}\,\{\lim_{i\to\infty} \nabla f(x_i) : x_i \to \bar{x}\}.$$

- This is counterintuitive since $\hat{\partial} f(\bar{x}) \subset \partial f(\bar{x}) \subset \partial_c f(\bar{x})$.

Example

$$f(x,y,z) = \begin{cases} \min\{x,y,z^2\} & ,\mathrm{if}\,(x,y,z) \in \mathbf{R}^3_+ \\ \min\{-x,-y,z^2\} & ,\mathrm{if}\,(x,y,z) \in \mathbf{R}^3_- \\ 0 & ,\mathrm{otherwise}. \end{cases}$$

Failure for the Clarke Subdifferential

- The theorem is false for the Clarke case!
- For Lipschitz f,

$$\partial_c f(\bar{x}) = \operatorname{conv}\{\lim_{i\to\infty} \nabla f(x_i) : x_i \to \bar{x}\}.$$

- This is counterintuitive since $\hat{\partial} f(\bar{x}) \subset \partial f(\bar{x}) \subset \partial_c f(\bar{x})$.

Example

$$f(x,y,z) = \begin{cases} \min\{x,y,z^2\} & ,\text{if}\,(x,y,z) \in \mathbf{R}^3_+ \\ \min\{-x,-y,z^2\} & ,\text{if}\,(x,y,z) \in \mathbf{R}^3_- \\ 0 & ,\text{otherwise.} \end{cases}$$

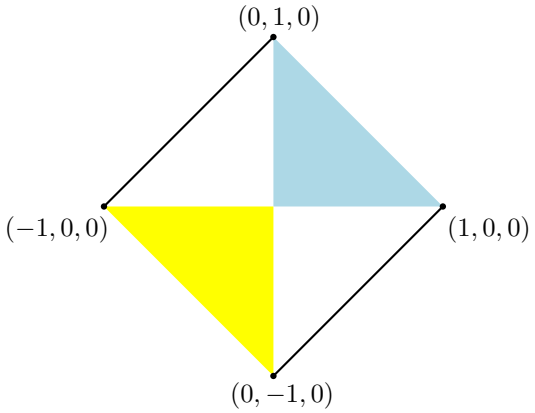


Figure: $\partial f(0,0,0)$

Work in progress

Duality: For a convex f,

$$M \text{ is identifiable at } \bar{x} \text{ for } \bar{v} \in \partial f(\bar{x})$$

iff

$$(\partial f)(M) \text{ is identifiable at } \bar{v} \text{ for } \bar{x} \in \partial f^*(\bar{v}).$$

Work in progress

Duality: For a convex f,

$$M \text{ is identifiable at } \bar{x} \text{ for } \bar{v} \in \partial f(\bar{x})$$

iff

$$(\partial f)(M) \text{ is identifiable at } \bar{v} \text{ for } \bar{x} \in \partial f^*(\bar{v}).$$

Minimality does not carry over.

Work in progress

Duality: For a convex f,

$$M \text{ is identifiable at } \bar{x} \text{ for } \bar{v} \in \partial f(\bar{x})$$

iff

$$(\partial f)(M) \text{ is identifiable at } \bar{v} \text{ for } \bar{x} \in \partial f^*(\bar{v}).$$

Minimality does not carry over.
However, often both M and $(\partial f)(M)$ are manifolds.

Work in progress

Duality: For a convex f,

$$M \text{ is identifiable at } \bar{x} \text{ for } \bar{v} \in \partial f(\bar{x})$$

iff

$$(\partial f)(M) \text{ is identifiable at } \bar{v} \text{ for } \bar{x} \in \partial f^*(\bar{v}).$$

Minimality does not carry over.
However, often both M and $(\partial f)(M)$ are manifolds.

Then have refined version of duality!

Work in progress

Duality: For a convex f,

$$M \text{ is identifiable at } \bar{x} \text{ for } \bar{v} \in \partial f(\bar{x})$$

iff

$$(\partial f)(M) \text{ is identifiable at } \bar{v} \text{ for } \bar{x} \in \partial f^*(\bar{v}).$$

Minimality does not carry over.
However, often both M and $(\partial f)(M)$ are manifolds.

Then have refined version of duality!

Example

Semi-definite cone $\mathbf{S}^n_+$ stratifies into "identifiable manifolds"

$$S_k := \{X \in \mathbf{S}^n_+ : \operatorname{rank} X = k\}, \ \textit{for } k = 0, \ldots, n,$$

and is self-dual. Then have rigorous duality between S_k and S_{n-k}.

Summary

- Presented the intuitive notion of identifiable sets.
- Showed how identifiable sets capture the essence of many previously developed concepts (critical cones, Partial Smoothness, optimality conditions).
- Generic existence.

Thank you.