Variational analysis with smooth substructure

Dmitriy Drusvyatskiy,
School of ORIE, Cornell University

May 22, 2012
Intuitive notion of identifiable sets.
Existence, calculus, properties.
Connection to critical cones (Generalized Reduction Lemma).
Illustration: Spectral functions.
Definition (Generalized critical points)

$\bar{x}$ is a critical point of $f : \mathbb{R}^n \to \mathbb{R}$ if $0 \in \partial f(\bar{x})$. 
Definition (Generalized critical points)
\( \bar{x} \) is a critical point of \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) if \( 0 \in \partial f(\bar{x}) \).

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

$$f_\nu(x) = f(x) - \langle \nu, x \rangle.$$  

[For simplicity],

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

Motivation (Sensitivity Analysis)
Consider the perturbed functions

\[ f_\nu(x) = f(x) - \langle \nu, x \rangle. \]

[For simplicity],

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

**Sensitivity question:** How do critical points of \( f_\nu \), near \( \bar{x} \), behave as \( \nu \) varies near \( \bar{\nu} \)?
Consider the perturbed functions

\[ f_\nu(x) = f(x) - \langle \nu, x \rangle. \quad \text{[For simplicity]}, \]

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

**Sensitivity question**: How do critical points of \( f_\nu \), near \( \bar{x} \), behave as \( \nu \) varies near \( \bar{\nu} \)? or equivalently how do solutions \( x_\nu \) of

\[ \nu \in \partial f(x), \]

vary, as we perturb \( \nu \) near \( \bar{\nu} \).
Motivating Example

Motivating example

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

Motivating example

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

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

Motivating Example

Motivating example

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

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

All perturbed solutions $x_\nu$ of $\nu \in \partial f(x)$ 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 \( f|_M \) matters!
Motivating Example

Motivating example

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

- 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.
Consider the system

\[ \nu \in \partial f(x). \]
Consider the system
\[ \nu \in \partial f(x). \]

**Definition (Identifiable sets)**

A set \( M \subset \mathbb{R}^n \) is identifiable at \( (\bar{x}, \bar{\nu}) \in \text{gph} \partial f \) if locally near \( \bar{x} \) have

\[ M = \pi_{\mathbb{R}^n}((U \times V) \cap \text{gph} \partial f), \]

for some neighbourhood \( U \times V \) of \( (\bar{x}, \bar{\nu}) \).
Consider the system
\[ \nu \in \partial f(x). \]

**Definition (Identifiable sets)**
A set \( M \subset \mathbb{R}^n \) is identifiable at \( (\bar{x}, \bar{v}) \in \text{gph} \partial f \) if locally near \( \bar{x} \) have
\[
M = \pi_{\mathbb{R}^n}((U \times V) \cap \text{gph} \partial f),
\]
for some neighbourhood \( U \times V \) of \( (\bar{x}, \bar{v}) \).

**Example (Trivial example)**

\[ f(x) = |x| \]

**Example (Normal cone map)**
Let \( \partial f = N_Q \) for a cube \( Q \subset \mathbb{R}^3 \). In this case \( M = \bar{x} + K_Q (\bar{x}, \bar{v}) \).
Finite identification

Consider the system

\[ \nu \in \partial f(x). \]

Definition (Identifiable sets)
A set \( M \subset \mathbb{R}^n \) is identifiable at \((\bar{x}, \bar{v}) \in \text{gph } \partial f\) if locally near \( \bar{x} \) have

\[ M = \pi_{\mathbb{R}^n}((U \times V) \cap \text{gph } \partial f), \]

for some neighbourhood \( U \times V \) of \((\bar{x}, \bar{v})\).

Example (Trivial example)
Finite identification

Consider the system

\[ v \in \partial f(x). \]

Definition (Identifiable sets)

A set \( M \subset \mathbb{R}^n \) is identifiable at \((\bar{x}, \bar{v}) \in gph \partial f\) if locally near \( \bar{x} \) have

\[ M = \pi_{\mathbb{R}^n}( (U \times V) \cap gph \partial f ), \]

for some neighbourhood \( U \times V \) of \((\bar{x}, \bar{v})\).

Example (Normal cone map)

Let \( \partial f = N_Q \) for a cube \( Q \subset \mathbb{R}^3 \).
Consider the system
\[ \nu \in \partial f(x). \]

**Definition (Identifiable sets)**
A set \( M \subset \mathbb{R}^n \) is identifiable at \( (\bar{x}, \bar{v}) \in \text{gph} \partial f \) if locally near \( \bar{x} \) have
\[ M = \pi_{\mathbb{R}^n}((U \times V) \cap \text{gph} \partial f), \]
for some neighbourhood \( U \times V \) of \( (\bar{x}, \bar{v}) \).

**Example (Normal cone map)**
Let \( \partial f = N_Q \) for a cube \( Q \subset \mathbb{R}^3 \).
Finite identification

Consider the system
\[ \nu \in \partial f(x). \]

**Definition (Identifiable sets)**
A set \( M \subset \mathbb{R}^n \) is identifiable at \( (\bar{x}, \bar{v}) \in \text{gph} \partial f \) if locally near \( \bar{x} \) have

\[ M = \pi_{\mathbb{R}^n}((U \times V) \cap \text{gph} \partial f), \]

for some neighbourhood \( U \times V \) of \( (\bar{x}, \bar{v}) \).

**Example (Normal cone map)**
Let \( \partial f = N_Q \) for a cube \( Q \subset \mathbb{R}^3 \).

\[ M = \bar{x} + K_Q(\bar{x}, \bar{v}). \]
Why are identifiable sets interesting?
Order of growth

Why are identifiable sets interesting?

Proposition (D, Lewis)

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

$\bar{x}$ 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 $\bar{x}$. 
Order of growth

Why are identifiable sets interesting?

Proposition (D, Lewis)

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

- $\bar{x}$ 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 $\bar{x}$. 
Locally minimal identifiable sets

Clearly all of $\mathbb{R}^n$ is identifiable at $(\bar{x}, \bar{y}) \in \text{gph} \, \partial f$ (not interesting). So...
Locally minimal identifiable sets

Clearly all of $\mathbb{R}^n$ is identifiable at $(\bar{x}, \bar{v}) \in \text{gph } \partial f$ (not interesting). So...

**Question:** What are the smallest possible identifiable sets?
Clearly all of $\mathbb{R}^n$ is identifiable at $(\bar{x}, \bar{v}) \in \text{gph } \partial f$ (not interesting). So...

**Question:** What are the smallest possible identifiable sets?

**Definition**

An identifiable set $M$ at $(\bar{x}, \bar{v}) \in \text{gph } \partial f$ is **locally minimal** if

$$M' \text{ identifiable at } (\bar{x}, \bar{v}) \rightarrow M \subset M', \text{ locally near } \bar{x}.$$
Locally minimal identifiable sets exist for

- fully amenable functions: $f(x) = g(F(x))$ where
  - $F$ is $C^2$-smooth,
  - $g$ is (convex) piecewise quadratic,
  - qualification condition holds.
Locally minimal identifiable sets exist for

- **fully amenable functions**: \( f(x) = g(F(x)) \) where
  - 1. \( F \) is \( C^2 \)-smooth,
  - 2. \( g \) is (convex) piecewise quadratic,
  - 3. qualification condition holds.

E.g. convex polyhedra, max-type functions, standard problems of nonlinear math programming.
Locally minimal identifiable sets exist for

- **fully amenable functions**: $f(x) = g(F(x))$ where
  - $F$ is $C^2$-smooth,
  - $g$ is (convex) piecewise quadratic,
  - qualification condition holds.

E.g. convex polyhedra, max-type functions, standard problems of nonlinear math programming.

A strong **chain rule** is available for composite functions

$$f(x) = g(F(x)).$$
The critical cone of a convex $Q$ at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$ is

$$K_Q(\bar{x}, \bar{v}) := T_Q(\bar{x}) \cap \bar{v}^\perp.$$
The critical cone of a convex $Q$ 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,
The critical cone of a convex $Q$ at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$ is

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

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
Proposition (Reduction Lemma due to Robinson)

If $Q$ is polyhedral, then

$$gph \, N_Q = gph \, N_{\bar{x} + K_Q(\bar{x}, \bar{\nu})} \text{ locally near } (\bar{x}, \bar{\nu}).$$
Proposition (Reduction Lemma due to Robinson)

If \( Q \) is polyhedral, then

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

Not true at all beyond polyhedral sets, but
**Dimension Reduction**

**Proposition (Reduction Lemma due to Robinson)**

*If* $Q$ *is polyhedral, then*

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

**Not true at all** beyond polyhedral sets, but

**Proposition (D, Lewis)**

*Let* $M$ *be a (prox-regular) identifiable set at* $(\bar{x}, \bar{v}) \in \text{gph } N_Q(\bar{x})$.

*Then*

$$\text{gph } N_Q = \text{gph } N_M \text{ locally near } (\bar{x}, \bar{v}),$$
Proposition (Reduction Lemma due to Robinson)

If $Q$ is polyhedral, then

\[ \text{gph } N_Q = \text{gph } N_{x + K_Q(x, v)} \text{ locally near } (x, v). \]

Not true at all beyond polyhedral sets, but

Proposition (D, Lewis)

Let $M$ be a (prox-regular) identifiable set at $(x, v) \in \text{gph } N_Q(x)$. Then

\[ \text{gph } N_Q = \text{gph } N_M \text{ locally near } (x, v), \]

and if $M$ is also locally minimal, then

\[ K_Q(x, v) = \text{cl conv } T_M(x). \]
Proposition (Reduction Lemma due to Robinson)

If $Q$ is polyhedral, then

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

Not true at all beyond polyhedral sets, but

Proposition (D, Lewis)

Let $M$ be a (prox-regular) identifiable set at $(\bar{x}, \bar{v}) \in \text{gph } N_Q(\bar{x})$. Then

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

and if $M$ is also locally minimal, then

$$K_Q(\bar{x}, \bar{v}) = \text{cl conv } T_M(\bar{x}).$$

May use this to study nonpolyhedral variational inequalities!
Identifiable manifolds

$M$ is an identifiable manifold at $(\bar{x}, \bar{v}) \in \text{gph} \partial f$ if $M$ is identifiable, $M$ is a manifold, and $f|_M$ is smooth.

Proposition (D-Lewis) Identifiable manifolds $M \subset \text{dom} f$ are automatically locally minimal. Identifiable manifolds provide a refinement of partly smooth manifolds introduced in Lewis’03.
Identifiable manifolds

$M$ is an identifiable manifold at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ if $M$ is identifiable, $M$ is a manifold, and $f\big|_M$ is smooth.
Identifiable manifolds

$M$ is an identifiable manifold at $(\bar{x}, \bar{v}) \in \text{gph } \partial f$ if $M$ is identifiable, $M$ is a manifold, and $f \big|_M$ is smooth.

**Proposition (D-Lewis)**

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

$M$ is an identifiable manifold at $(\bar{x}, \bar{v}) \in \text{gph} \, \partial f$ if $M$ is identifiable, $M$ is a manifold, and $f\big|_{M}$ is smooth.

Proposition (D-Lewis)

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

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

$M$ is an identifiable manifold at $(\tilde{x}, \tilde{v}) \in \text{gph} \, \partial f$ if $M$ is identifiable, $M$ is a manifold, and $f\big|_M$ is smooth.

Proposition (D-Lewis)

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

- Identifiable manifolds provide a refinement of partly smooth manifolds introduced in Lewis ’03.
When an identifiable manifold exists, nonsmoothness is not intrinsic.
Identifiable manifolds

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

Consider $S^n := \{ n \times n \text{ symmetric matrices} \}$ and the eigenvalue map

$$A \mapsto (\lambda_1(A), \ldots, \lambda_n(A)),$$

where

$$\lambda_1(A) \leq \ldots \leq \lambda_n(A).$$
Consider $\mathbb{S}^n := \{n \times n$ symmetric matrices$\}$ and the eigenvalue map

$$A \mapsto (\lambda_1(A), \ldots, \lambda_n(A)),$$

where

$$\lambda_1(A) \leq \ldots \leq \lambda_n(A).$$

For $f : \mathbb{R}^n \rightarrow \mathbb{R}$, invariant under permutation of coordinates, form the spectral function

$$f \circ \lambda : \mathbb{S}^n \rightarrow \mathbb{R}.$$
Lifts of identifiable manifolds

Consider $S^n := \{n \times n$ symmetric matrices$\}$ and the eigenvalue map

$$A \mapsto (\lambda_1(A), \ldots, \lambda_n(A)),$$

where

$$\lambda_1(A) \leq \ldots \leq \lambda_n(A).$$

For $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, invariant under permutation of coordinates, form the spectral function

$$f \circ \lambda : S^n \to \overline{\mathbb{R}}.$$

(e.g. $(f \circ \lambda)(A) = \lambda_n(A)$ or $(f \circ \lambda)(A) = \sum_i |\lambda_i(A)|$).
Lifts of identifiable manifolds

Consider $\mathbf{S}^n := \{n \times n$ symmetric matrices$\}$ and the eigenvalue map

$$A \mapsto (\lambda_1(A), \ldots, \lambda_n(A)),$$

where

$$\lambda_1(A) \leq \ldots \leq \lambda_n(A).$$

For $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, invariant under permutation of coordinates, form the spectral function

$$f \circ \lambda : \mathbf{S}^n \to \overline{\mathbb{R}}.$$

(e.g. $(f \circ \lambda)(A) = \lambda_n(A)$ or $(f \circ \lambda)(A) = \sum_i |\lambda_i(A)|$).

Identifiable manifolds “lift”: (D, Lewis), (Daniilidis, Malick, Sendov)

$\mathcal{M}$ identifiable manifold at $(\bar{x}, \bar{v}) \in \text{gph } \partial f$

$\implies \lambda^{-1}(\mathcal{M})$ identifiable manifold at $(\bar{X}, \bar{V}) \in \text{gph } \partial (f \circ \lambda)$. 
Study “facial” structure of spectral sets (and functions). E.g.

\[ S_+^n = \bigcup_{k=1}^{n} \{ X \in S_+^n : \text{rank } X = k \}. \]
Study “facial” structure of spectral sets (and functions). E.g.

$$S^n_+ = \bigcup_{k=1}^n \{ X \in S^n_+ : \text{rank } X = k \}.$$ 

May lead to sensitivity analysis, provided can project onto $M$ easily.
• Presented the intuitive notion of **identifiable sets**.
• Showed how identifiable sets capture the essence of previously developed concepts (**dimension reduction, critical cones, optimality conditions**).
• Application to **spectral functions**.
Thank you.