Variational analysis with smooth substructure

Dmitriy Drusvyatskiy,
School of ORIE, Cornell University

May 22, 2012

Goals

- Intuitive notion of identifiable sets.
- Existence, calculus, properties.
- Connection to critical cones (Generalized Reduction Lemma).
- Illustration: Spectral functions.

Critical points

Definition (Generalized critical points)

$\bar{x}$ is a critical point of $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ if $0 \in \partial f(\bar{x})$.

Critical points

Definition (Generalized critical points)

$\bar{x}$ is a critical point of $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ 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 $\bar{v}$? or equivalently how do solutions x_v of

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

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

Motivating example

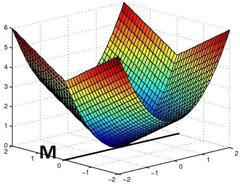


Figure: $f(x, y) = x^2 + |y|, \quad M = \{(t, 0) : -1 < t < 1\}$

Motivating example

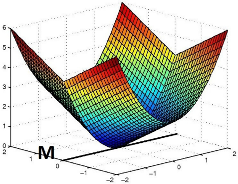


Figure: $f(x,y) = x^2 + |y|, \quad M = \{(t,0) : -1 < t < 1\}$

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

Motivating Example

Motivating example

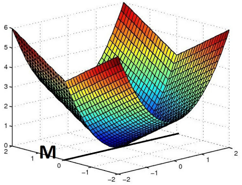


Figure: $f(x,y) = x^2 + |y|, \quad M = \{(t,0) : -1 < t < 1\}$

- 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

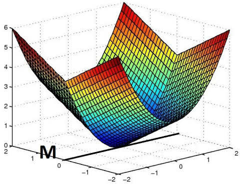


Figure: $f(x, y) = x^2 + |y|, \quad M = \{(t, 0) : -1 < t < 1\}$

- 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

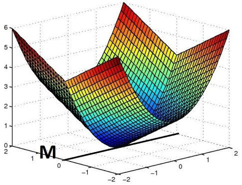


Figure: $f(x,y) = x^2 + |y|, \quad M = \{(t,0) : -1 < t < 1\}$

- 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 $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ if locally near $\bar{x}$ have

$$M = \pi_{\mathbf{R}^n}((U \times V) \cap \operatorname{gph} \partial f),$$

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

Finite identification

Consider the system

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

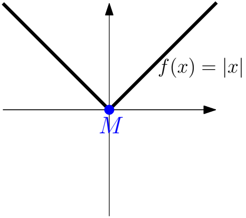
Definition (Identifiable sets)

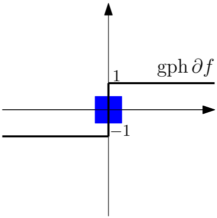
A set $M \subset \mathbf{R}^n$ is identifiable at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ if locally near $\bar{x}$ have

$$M = \pi_{\mathbf{R}^n}((U \times V) \cap \operatorname{gph} \partial f),$$

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

Example (Trivial example)





Finite identification

Consider the system

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

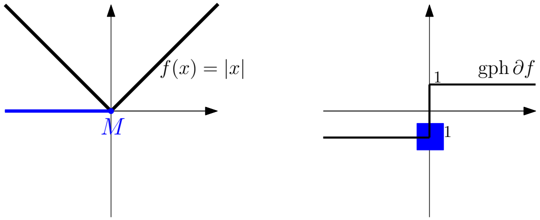
Definition (Identifiable sets)

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

$$M = \pi_{\mathbf{R}^n}((U \times V) \cap \operatorname{gph} \partial f),$$

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)

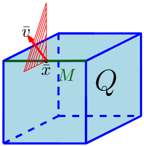
A set $M \subset \mathbf{R}^n$ is identifiable at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ if locally near $\bar{x}$ have

$$M = \pi_{\mathbf{R}^n}((U \times V) \cap \operatorname{gph} \partial f),$$

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

Example (Normal cone map)

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



Finite identification

Consider the system

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

Definition (Identifiable sets)

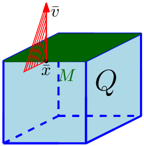
A set $M \subset \mathbf{R}^n$ is identifiable at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ if locally near $\bar{x}$ have

$$M = \pi_{\mathbf{R}^n}((U \times V) \cap \operatorname{gph} \partial f),$$

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

Example (Normal cone map)

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



Finite identification

Consider the system

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

Definition (Identifiable sets)

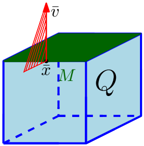
A set $M \subset \mathbf{R}^n$ is identifiable at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ if locally near $\bar{x}$ have

$$M = \pi_{\mathbf{R}^n}((U \times V) \cap \operatorname{gph} \partial f),$$

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

Example (Normal cone map)

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



In this case $M = \bar{x} + K_Q(\bar{x}, \bar{v})$.

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 $(\bar{x}, 0) \in \operatorname{gph} \partial f$.

Order of growth

Why are identifiable sets interesting?

Proposition (D, Lewis)

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

- *$\bar{x}$ 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 $\mathbf{R}^n$ is identifiable at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ (not interesting).
So...

Locally minimal identifiable sets

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

Locally minimal identifiable sets

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

Definition

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

$$M' \text{ identifiable at } (\bar{x}, \bar{v}) \implies M \subset M', \text{ locally near } \bar{x}.$$

Existence and calculus

Locally minimal identifiable sets exist for

- 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

- 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.

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

Existence and calculus

Locally minimal identifiable sets exist for

- 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.

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

A strong chain rule is available for composite functions

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

Dimension Reduction

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

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

Dimension Reduction

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,

Dimension Reduction

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, because of

Dimension Reduction

Proposition (Reduction Lemma due to Robinson)

If Q is polyhedral, then

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

Dimension Reduction

Proposition (Reduction Lemma due to Robinson)

If Q is polyhedral, then

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

Not true at all beyond polyhedral sets, but

Dimension Reduction

Proposition (Reduction Lemma due to Robinson)

If Q is polyhedral, then

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

Not true at all beyond polyhedral sets, but

Proposition (D, Lewis)

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

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

Dimension Reduction

Proposition (Reduction Lemma due to Robinson)

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) identifiable set at $(\bar{x},\bar{v}) \in \operatorname{gph} N_Q(\bar{x})$. Then

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

and if M is also locally minimal, then

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

Dimension Reduction

Proposition (Reduction Lemma due to Robinson)

If Q is polyhedral, then

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

Not true at all beyond polyhedral sets, but

Proposition (D, Lewis)

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

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

and if M is also locally minimal, then

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

May use this to study nonpolyhedral variational inequalities!

Identifiable manifolds

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

Proposition (D-Lewis)

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

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.

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.

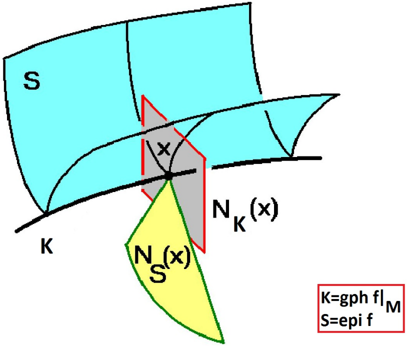
Identifiable manifolds

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.

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

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

where

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

Lifts of identifiable manifolds

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

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

where

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

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

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

Lifts of identifiable manifolds

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

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

where

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

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

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

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

Lifts of identifiable manifolds

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

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

where

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

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

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

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

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

$$M \text{ identifiable manifold at } (\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$$
$$\implies \lambda^{-1}(M) \text{ identifiable manifold at } (\bar{X}, \bar{V}) \in \operatorname{gph} \partial (f \circ \lambda).$$

Lifts of identifiable manifolds

Study "facial" structure of spectral sets (and functions). E.g.

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

Lifts of identifiable manifolds

Study "facial" structure of spectral sets (and functions). E.g.

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

May lead to sensitivity analysis, provided can project onto M easily.

Summary

- 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.