Lecture 17: Clarke & his differential.

Announcement: \( \text{This is \#2 tomorrow \& last week.} \)

- Why Clarke differential?
  - Need a model of a function.
  - Doesn't involve python; hard to predict what is used in 10 years.

Subgradients:

We saw it in optimization:

\[
\partial_f(w) = \text{"set of lower bounding slopes at } w \" = \{ \alpha \in \mathbb{R}^d : \forall v, f(v) \geq f(w) + \langle \alpha, v-w \rangle \}.
\]

Example 1:

\[
f(v) = \langle v, w \rangle + 1 \]

\[
\partial f(w) = \mathbb{R}^d
\]

Example 2:

\[
f(v) = \|v\|_1\]}

\[
\partial f(v) = \text{sign}(v)
\]

Example 3:

\[
f(v) = \text{max}(v)
\]

\[
\partial f(w) = \{ x \in \mathbb{R}^d : x \geq 0, x \leq w \}
\]

Properties (for convex \( f: \mathbb{R}^d \rightarrow \mathbb{R} \)):

\( \text{If } \partial f \text{ is nonempty everywhere, } \partial f \text{ is closed convex.} \)

\( \text{If } \partial f \text{ is directionally differentiable, } \)

\[
f'(w, v) = \lim_{h \to 0} \frac{f(w+hv)-f(w)}{h} = \sup_{x \in \partial f(w)} \langle v, x \rangle
\]

\( \text{Mean value theorem; see for instance} \)

- Hiriart-Urruty & Lemaréchal “Fundamentals of convex analysis”
  - (many pictures)
- Rockafellar & Wets “Convex analysis” and linear optimization.

\( \text{We used it in our first set proof.} \)

\( \text{Proposition (Rock's inequality). If } f(\text{max}(x, y)) \text{ and } X \text{ is a uv.} \)

\[
\text{then } f(\text{max}(x, y)) = f(x) + \langle y - x, \partial f(x) \rangle
\]

Proof: \( \forall x \in \partial f(x), \langle v, x \rangle \geq f(x) + \langle v, x - bx \rangle \)

\[= f(x) + \langle y - x, \partial f(x) \rangle. \]


Clarke Dependent

First definition of Clarke: The function decides the function makes it
Recall: Direct function $f(x)$ is $\frac{\partial}{\partial x} f(x) = \frac{1}{2} \frac{\partial}{\partial x} g(x)$

Then $g(x) = e^x$: Yes, $\frac{\partial}{\partial x} g(x) = (e^x)^2$

Define: $g(x) = \frac{1}{2} \frac{\partial}{\partial x} g(x)$

$g(x) = \left( \frac{1}{2} \frac{\partial}{\partial x} g(x) \right)^2$ is called $x^2$...

Results: Consider the equation with Clarke dependence.

What we want:

$0$ for one variable
$0$ for one function
$0$ for multiple variables
$0$ for multiple functions
$0$ for randomly chosen functions
$0$ for randomly chosen

Recall:  $1 + 1$ is locally Lipschitz

Prop.: $\exists x_0$ s.t. $f(x)$ is locally Lipschitz

Proof: $\exists x_0$ s.t. $f(x)$ is locally Lipschitz

Theorem: $\exists x_0$ s.t. $f(x)$ is locally Lipschitz

Remarks: $\exists x_0$ s.t. $f(x)$ is locally Lipschitz

$x + e^x$