Lecture 21

Last time

- P-D guarantee continued
- Frechet subdifferential
- Density Theorem

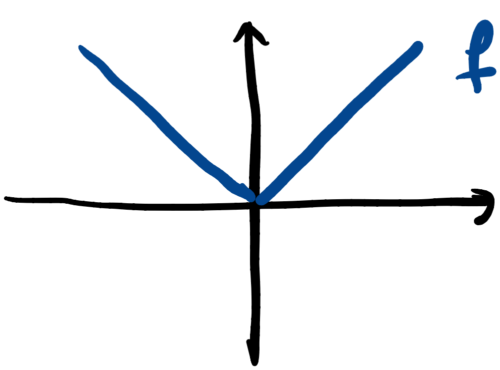
Today

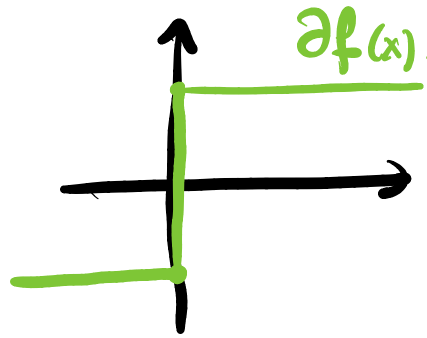
- Fuzzy Calculus

Fuzzy Calculus

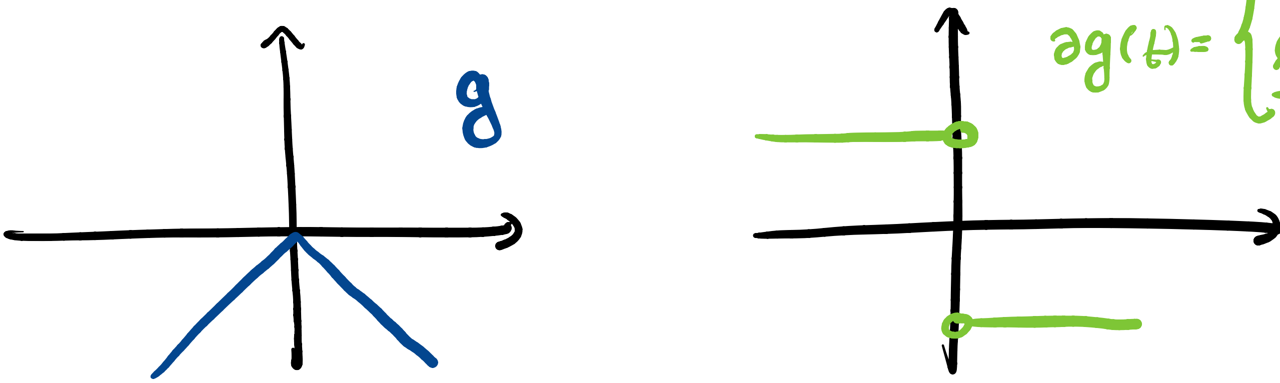
The properties of the Frechet subdifferential are weaker than what we had for convex subdifferentials. How about calculus rules? Well, it turns out that they can also fail spectacularly.

Example: Consider two simple functions $f:\mathbb{R}\to\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$ given by $f(t)=|t|$ and $g(t)=-|t|$.





$$\partial f(x) = \begin{cases} -1 & t<0 \\ [-1,1] & t=0 \\ 1 & t>0 \end{cases}$$



$$\partial g(t) = \begin{cases} 1 & t<0 \\ \emptyset & t=0 \\ -1 & t>0 \end{cases}$$

Then $f+g \equiv 0$ and

$\partial(f+g)(0) = \{0\}$ and $\partial f(0) + \partial g(0) = \emptyset$.

However, there are "fuzzy" calculus rules.

Theorem (Fuzzy chain rule): Consider the function $f(x) = h(c(x))$ where $h: Y \to \mathbb{R} \cup \{\infty\}$ is a proper, closed function and $c: E \to Y$ is differentiable around a point $\bar{x} \in E$.
Then, the inclusion holds:

$$(\supseteq) \qquad \partial f(\bar{x}) \supseteq \nabla c(\bar{x})^* \partial h(c(\bar{x})).$$

Conversely, for every subgradient $v \in \partial f(x)$ and $\varepsilon > 0$, there exists points $x \in E$ and $y \in Y$ such that

$$\|x - \bar{x}\| \le \varepsilon, \quad \|y - c(\bar{x})\| \le \varepsilon, \quad |h(y) - h(c(\bar{x}))| \le \varepsilon,$$

and

$$(\subseteq) \qquad v \in \nabla c(x)^* \partial h(y) + \varepsilon B.$$

Before we proceed this result we note that it implies a sum rule thanks to the following separable sum rule.

Lemma (separable sum rule)
Let $f : E^k \to \mathbb{R} \cup \{+\infty\}$ be a closed proper function s.t $f(x_1, \dots, x_k) = \sum_{i=1}^{k} f_i(x_i)$ with f_i closed, proper. Then,

$$\partial f(x_1, \dots, x_k) = \begin{bmatrix} \partial f_1(x_1) \\ \vdots \\ \partial f_k(x_k) \end{bmatrix}. \quad +$$

Theorem (Fuzzy sum rule): Consider a pair of closed, proper functions $f_1, f_2 : E \to \mathbb{R} \cup \{\infty\}$. Then,

$$\partial (f_1 + f_2)(x) \supseteq \partial f_1(x) + \partial f_2(x) \quad \forall x.$$

Moreover, for any $v \in \partial(f_1 + f_2)(x)$ and $\varepsilon > 0$, we have $\exists\, x_1, x_2 \in x + \varepsilon B$ s.t.

$$v \in \partial f_1(x_1) + \partial f_2(x_2) + \varepsilon B.$$

Proof: Apply the fuzzy chain rule with $f(x_1, x_2) = f_1(x_1) + f_2(x_2)$ and

$c(x) = \binom{x}{x}$. Then, $\exists\, y=(x_1,x_2) \in c(x) + \varepsilon B$ and $\nabla c(x)^* = [I \;\; I]$ and so

$$x_1, x_2 \in x + \varepsilon B$$

and $$v \in \nabla c(x)^* \partial f(y) + \varepsilon B$$

Separable sum rule
$$= [I \;\; I] \begin{bmatrix} \partial f_1(x_1) \\ \partial f_2(x_2) \end{bmatrix} + \varepsilon B$$

$$= \partial f_1(x_1) + \partial f_2(x_2) + \varepsilon B.$$

$\square$

Proof of Fuzzy Chain Rule: Before we state a nontrivial fact about Frechet subdifferentials.

Claim (Viscosity subgradients): Consider a function $f : E \to \mathbb{R} \cup \{+\infty\}$ and $x \in \operatorname{dom} f$. Then, $v \in \partial f(x)$ if, and only if, there exists a C^1-function $w : U \to \mathbb{R}$ defined on a neighborhood U of x satisfying $w(x) = f(x)$, $\nabla w(x) = v$, and $w(y) < f(y)$ for all $y \in U \setminus \{x\}$. $\dashv$

We leave this claim as an exercise.

Inclusion $(\supseteq)$ is easy to prove using simply the definitions of subdifferentials and Jacobians. We focus on $(\subseteq)$. Let $v \in \partial f(\bar{x})$. By the claim above $\exists\, \omega : U \to \mathbb{R}$ a C^1-function with $\nabla \omega(\bar{x}) = v$, $f(\bar{x}) = \omega(\bar{x})$, and $f(x) > \omega(\bar{x})$ $\forall x \in U \setminus \{\bar{x}\}$. Thus, $\bar{x}$ is the unique minimizer of

$$\min_{x \in U} \; f(x) - \omega(x)$$

with $f(\bar{x}) - \omega(\bar{x}) = 0$. Shrinking U, we might assume it is closed and c is differentiable on U. Let $V \subseteq Y$ be a closed ball around $c(\bar{x})$ such that $h(y) \geq h(c(\bar{x})) - 1$ $\forall y \in V$. Consider the sequence of problems

$$(\bowtie) \quad \min_{x \in U,\, y \in V} \; \underbrace{h(y) - \omega(x) + \frac{n}{2}\|y - c(x)\|^2}_{F_n(x,y)}$$

Since these losses are closed and coercive, then there exists a sequence of minimizers (x_n, y_n). Let us show this sequence converges to $(\bar{x}, c(\bar{x}))$. Since the sequence is bounded, we might assume WLOG $(x_n, c(x_n)) \to (x^*, y^*)$. Note that

$$F_n(x_n, y_n) \leq F(\bar{x}, c(\bar{x})) = 0.$$

Thus,

$$\frac{n}{2}\|y_n - c(x_n)\|^2 \leq \omega(x_n) - h(y_n).$$

the right-hand is bounded since ω is C^1 and $-h(y_n) \leq -h(c(\bar{x})) + 1 \quad \forall n$. Then, dividing by n and taking limits yields $y^* = c(x^*)$.

Closedness ensures

$$\begin{aligned}
f(x^*) - \omega(x^*) &= h(y^*) - \omega(x^*) \\
(\to) \quad &\leq \liminf_{n\to\infty} h(y_n) - \omega(x_n) - \frac{n}{2}\|y_n - c(x_n)\|^2 \\
&= \liminf_{n\to\infty} F_n(x_n, y_n)
\end{aligned}$$

$$\leq \liminf_{n\to\infty} F_n(\bar{x}, c(\bar{x})) = \underset{= \min_{x\in U} f(x)-\omega(x)}{0}.$$

Hence, x^* minimizes $f-\omega$ over U. Since $\bar{x}$ was the unique minimizer of this problem, $x^* = \bar{x}$. Then, equality holds throughout and so $h(y_n) \to h(y^*)$. Since $x_n \to \bar{x}$, then for large n, $x_n \in \operatorname{int} U$, optimality conditions for $(*)$ read

$$0 \in -\nabla\omega(x_n) - n\nabla c(x)^*(y_n - c(x_n))$$

$$0 \in \partial h(y_n) + n(y_n - c(x_n))$$

Substituting the second inclusion in the first one and adding and substracting v on both sides yield

$$v \in \nabla c(x)^* \partial h(y_n) + \underbrace{v - \nabla\omega(x_n)}_{\to 0 \text{ as } n\to\infty}$$

The theorem follows immediately. $\square$