Supplement: classic convex/geometry theorems

Tuesday, January 19, 2021 10:34 PM

Not needed for the course

Rademacher's Thm $\mathcal{U} \subseteq \mathbb{R}^n$ open, $f: \mathcal{U} \to \mathbb{R}^m$ Lipschitz continuous (i.e. Hölder w/ exponent 1)

then f is differentiable almost everywhere (a.e.)

(not true if we weaken, eg. Hölder w/ exponent in $(0,1)$.

There are cts, nowhere differentiable fcn like $\sum_{n=1}^{\infty} a^n \cos(b^n x)$ for certain a, b)

Alexandrov's Thm $\mathcal{U} \subseteq \mathbb{R}^n$ open, $f: \mathcal{U} \to \mathbb{R}$ convex

then $\nabla^2 f$ exists a.e.

⚠ Differentiable a.e. doesn't mean we can "pretend" it's differentiable.

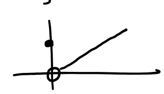
eg, minimizers may be at these non-diff. pts, like $x=0$ for $\min |x|$

NLP Myth #13 (Harvey Greenberg's collection of myths)

"A convex fcn is cts." FALSE

True on the relative interior of its effective domain, not on boundary

Ex: $f: \mathbb{R}_+ \to \mathbb{R}$, $f(x) = \begin{cases} 1 & x=0 \\ x & x \neq 0 \end{cases}$



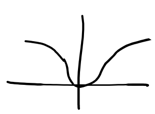
NLP Myth #1, modified $f \in C^1 \Rightarrow$ 1st order Taylor series will converge to $f(x)$

FALSE: it'll converge but not nec. to right #

(need f to be analytic)

Ex: $f(x) = \begin{cases} 0 & x=0 \\ e^{-1/x^2} & x \neq 0 \end{cases}$ $f \in C^\infty$ but not analytic

$f(0) = f'(0) = f''(0) = \ldots = 0$



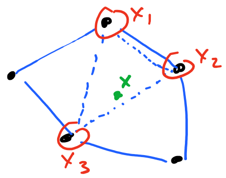
Myth $f(x,y)$ cts in x, cts in y $\Rightarrow$ f is cts (i.e. "jointly cts")

False. wikipedia has counterexamples

Geometric Thms

Thm: Carathéodory $x \in \operatorname{conv}(C) \subseteq \mathbb{R}^n$ then $\exists$ $n+1$ pts

$\{x_1, x_2, \ldots, x_{n+1}\} \subseteq C$ s.t. $x \in \operatorname{conv}(\{x_1, x_2, \ldots, x_{n+1}\})$.



$\mathbb{R}^2$ so $n=2$ $\qquad C =$ black dots •

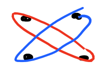
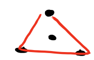
proof via basic linear algebra

Thm: Helly 1913 $\{A_1, \ldots, A_m\} \subseteq \mathbb{R}^n$, all A_i convex, $m > n$

if the intersection of all subsets of size $n+1$ are non-empty,

then $\bigcap_{j=1}^{m} A_j \neq \emptyset$.

Radon's Thm 1921 Any set of $n+2$ pts. in $\mathbb{R}^n$ can be partitioned into 2 disjoint subsets whose convex hulls intersect

Ex: $\mathbb{R}^2$: $4 = n+2$ pts  or 

3 pts, not possible always 

Krein-Milman / Minkowski 1940 "only need corners if convex"

Let C be a convex and compact set, then

$C = \overline{\text{conv}}(\text{ext}(C))$,

$\text{ext}(C)$ = extreme points (generalizes a "vertex" of a polytope)

$x \in \text{ext}(C)$ if $x \in C$ and x isn't contained in any open line segment joining points in C

C  ext(C)

Differentiability in $\mathbb{R}^n$ $n \geq 1$

There are different notions of differentiability. For $\mathbb{R}^1$, these all coincide luckily

1) (weakest) Partial derivatives exist, i.e., directional derivatives along coordinate axes

i.e., $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$ all exist

ex, $\mathbb{R}^2$, $f(x,y) = (xy)^{1/3}$, $\frac{\partial f}{\partial x} = \frac{1}{3} x^{-2/3} y^{1/3}$ and $\frac{\partial f}{\partial y}$ also exists

but along line $y=x$, let $g(x) = f(x,x) = x^{2/3}$, not differentiable at 0 since $g'(x) = \frac{2}{3} x^{-1/3}$

2) (next weakest) Gateaux differentiable, i.e., directional derivatives exist for all directions
version 1

i.e., $\forall$ directions $d \in \mathbb{R}^n$, $f'(x; d) := \lim_{h \to 0} \frac{f(x+h \cdot d) - f(x)}{h}$ exists.

2′) (next weeks) Gateaux diff, version 2 (authors don't agree)

same as 2) but also require $d \mapsto f'(x;d)$ is a bounded linear function "$\nabla f(x)$"

in $\mathbb{R}^n$, this comes for free

Saying it's linear means, in a Hilbert Space (i.e., using Riesz R)

we can write $f'(x;d) = \boxed{\langle \nabla f(x), d \rangle}$ COMMON NOTATION

3) (strictest) Fréchet differentiable

means $d \mapsto f'(x;d)$ is a linear function (like 2′)

and there's a uniform rate of convergence (in "h") independent of the direction,

i.e., $$\lim_{\|d\| \to 0} \frac{\left\| \left(f(x) + \langle \nabla f(x), d \rangle \right) - f(x+d) \right\|}{\|d\|} = 0$$

in case $f: \mathbb{R}^n \to \mathbb{R}^m$, $m > 1$

4) (even stricter than strict)

$f \in C^1$ i.e., $\nabla f(x)$ exists $\forall x$ and it's continuous

This implies Fréchet (hence Gateaux) diff. *

[So for simplicity, we usually assume $f \in C^1$ and don't worry about the details]

* I'm pretty sure but not 100% -- it's not obvious.

(In particular, we often assume ∇f is Lipschitz continuous, even stronger assumption than $f \in C^1$!)