HW #5 due next Tues.

Convex functions, Part II

Lecture 14

November 4, 2021

Beck, section 7.4-7.6

Recall (from last week)

f convex on C $\iff$ f satisfies "gradient inequality"

$$f(\underline{y}) - f(\underline{x}) \geq \nabla f(\underline{x})^T(\underline{y} - \underline{x}) \quad \forall\, \underline{x}, \underline{y} \in C$$

$$\left.\begin{array}{c} \nabla f(\underline{x}) = 0 \\ + \\ f \text{ convex} \end{array}\right\} \Longrightarrow \underline{x} \text{ is a global minimizer.}$$

f convex on C $\iff$ $\nabla^2 f(\underline{x}) \succeq \underline{\underline{0}} \quad \forall\, \underline{x} \in C$

f quadratic:
$$\nabla^2 f \succeq \underline{\underline{0}} \iff f \text{ convex}$$
$$\nabla^2 f \succ \underline{\underline{0}} \iff f \text{ strictly convex.}$$

Convexity-preserving operations

Goal: collect properties of convex functions that make identifying convex functions easier.

First property:

<u>Proposition</u>: Let $f_1, f_2 \ldots f_k$ be convex functions mapping a convex set $C \subset \mathbb{R}^n$ to $\mathbb{R}$. Let $\underline{\lambda} \in \mathbb{R}^k_+$. Then

$$\sum_{j=1}^{k} \lambda_j f_j(\underline{x})$$ is a convex function.

(Conic combinations of convex functions are convex)

Proof: Define $g(\underline{x}) = \sum_{j=1}^{k} \lambda_j f_j(\underline{x})$

Let $t \in (0,1)$:

$$g(t\underline{x} + (1-t)\underline{y}) = \sum_{j=1}^{k} \lambda_j f_j(t\underline{x} + (1-t)\underline{y})$$

convexity of f_j

$$\leq \sum_{j=1}^{k} \lambda_j [t f_j(\underline{x}) + (1-t) f_j(\underline{y})]$$

$$= t \sum_{j=1}^{k} \lambda_j f_j(\underline{x}) + (1-t) \sum_{j=1}^{k} \lambda_j f_j(\underline{y})$$

$$= t g(\underline{x}) + (1-t) g(\underline{y}).$$ ∎

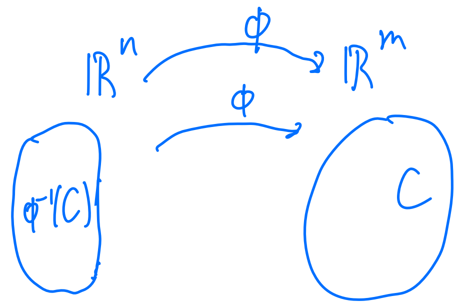
Compositions w/ affine functions.

Definition: $\phi : \mathbb{R}^n \to \mathbb{R}^m$ is <u>affine</u> if $\phi(\underline{x}) = \underline{\underline{A}}\,\underline{x} + \underline{b}$,

$\underline{\underline{A}} \in \mathbb{R}^{n \times m}$, $\underline{b} \in \mathbb{R}^m$.

Lemma: Let $\phi : \mathbb{R}^n \to \mathbb{R}^m$ be affine, and let $C \subset \mathbb{R}^m$ be a convex set. Then

$$\phi^{-1}(C) = \{ \underline{x} \in \mathbb{R}^n \mid \phi(\underline{x}) \in C \}$$ is convex in $\mathbb{R}^n$



Proof sketch:

$$\underline{x}, \underline{y} \in \phi^{-1}(C)$$

$$\Downarrow \ (\phi \text{ def'n})$$

$$\phi(\underline{x}), \ \phi(\underline{y}) \in C$$

$$\Downarrow \ (C \text{ convex})$$

$$\lambda \phi(\underline{x}) + (1-\lambda)\phi(\underline{y}) \in C$$

$$\Downarrow \ \phi(\underline{x}) = \underline{\underline{A}}\,\underline{x} + \underline{b}$$

$$\lambda(\underline{\underline{A}}\,\underline{x} + \underline{b}) + (1-\lambda)(\underline{\underline{A}}\,\underline{y} + \underline{b}) \in C$$

$$\underline{\underline{A}}[\lambda \underline{x} + (1-\lambda)\underline{y}] + \underline{b} \in C$$

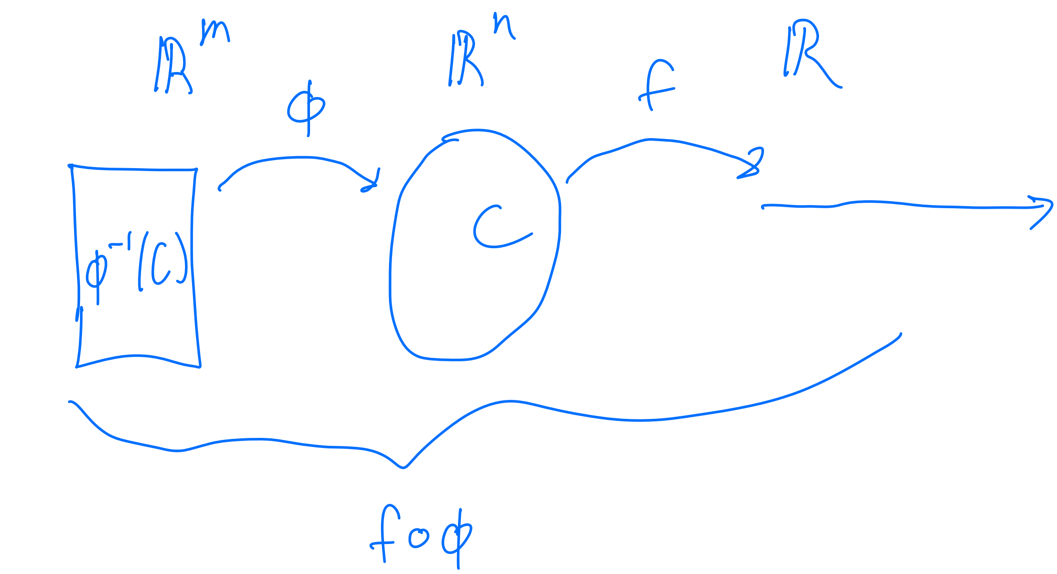
$$\phi(\lambda \underline{x} + (1-\lambda)\underline{y}) \in C$$

$$\Downarrow \ \phi \text{ def'n}$$

$$\lambda \underline{x} + (1-\lambda)\underline{y} \in \phi^{-1}(C) \ \blacksquare$$

Proposition: Let $f: C \to \mathbb{R}$ be convex, and let $\phi: \mathbb{R}^m \to \mathbb{R}^n$ be affine.

Then $f(\phi) = f \circ \phi$ is convex over the convex set $\phi^{-1}(C)$.



<u>Proof</u>: Use Defn...

<u>Example</u>: Show $f(\underline{x}) = \|\underline{\underline{A}}\,\underline{x}\|_2$ is convex

Note: $g(\underline{y}) = \|\underline{y}\|_2$, $\underline{y} \in \mathbb{R}^m$ is convex

$\phi(\underline{x}) = \underline{\underline{A}}\,\underline{x}$ is affine

composition property

$\Rightarrow f(\underline{x}) = \|\underline{\underline{A}}\,\underline{x}\|_2 = g(\phi(\underline{x})) \Rightarrow f$ convex.

Example: Show $f(\underline{x}) = \sqrt{\|\underline{x}\|_2^2 + 1}$ is convex.

(wait until later...)

Note: ϕ affine, f convex $\not\Rightarrow$ $\phi \circ f$ is convex.

Examples

Example (Beck, 7.19)

Show that

$$f(x_1, x_2) = x_1^2 + 2x_1x_2 + 3x_2^2 + 2x_1 - 3x_2 + e^{x_1},$$

is convex.

$f(\underline{x}) = g(\underline{x}) + h(\underline{x})$

$g(\underline{x}) = x_1^2 + 2x_1x_2 + 3x_2^2 + 2x_1 - 3x_2$

$h(\underline{x}) = e^{x_1}$

h is convex (e.g compute Hessian...)

g is convex: $g(\underline{x}) = \underline{x}^T \underline{\underline{A}}\, \underline{x} + 2\underline{b}^T \underline{x}$

$$\underline{\underline{A}} = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}, \quad \underline{b} = \begin{pmatrix} 1 \\ -3/2 \end{pmatrix}$$

$\underline{\underline{A}} > \underline{\underline{0}}$ since $\det(\underline{\underline{A}}) = 2 > 0$ $\Longrightarrow \lambda_1, \lambda_2 > 0.$

$\mathrm{Tr}(\underline{\underline{A}}) = 4 > 0$

$\Longrightarrow f = g + h$ is convex.

Examples

Example (Beck, 7.20)

Show that

$$f(x_1, x_2, x_3) = e^{x_1 - x_2 + x_3} + e^{2x_2} + x_1,$$

is convex.

$f(\underline{x}) = g_1(\underline{x}) + g_2(\underline{x}) + g_3(\underline{x})$

$g_1(\underline{x}) = e^{x_1 - x_2 + x_3} = h(\phi(\underline{x})),\quad h(x) = e^x$

$\Rightarrow g_1$ is convex. $\quad \phi(\underline{x}) = x_1 - x_2 + x_3$

$g_2(\underline{x}) = e^{2x_2} = \tilde{h}(\tilde{\varphi}(\underline{x})),\quad \tilde{h}(x) = e^x$

$\tilde{\varphi}(\underline{x}) = 2x_2$

$\Rightarrow g_2$ is convex.

$g_3(\underline{x}) = x_1 \Rightarrow g_3$ is convex.

$\Rightarrow f = g_1 + g_2 + g_3$ is convex.

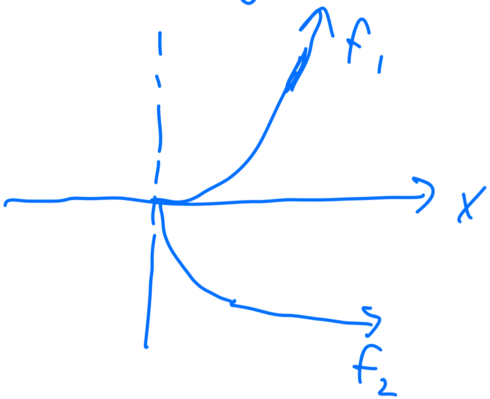
Composition of general convex functions

f_1 and f_2 are convex.

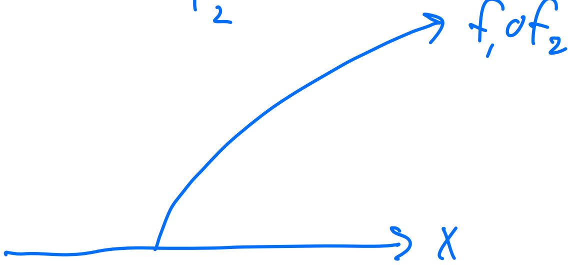
$f_1 \circ f_2 = f_1(f_2)$ is not necessarily convex.

E.g.: $f_1(x) = x^2$

$f_2(x) = -x^{1/3}$



$$(f_1 \circ f_2)(x) = \left(-x^{1/3}\right)^2 = x^{2/3}$$



Monotonic functions

Def'n: $f : \mathbb{R} \to \mathbb{R}$ is monotone increasing if

$$x \le y \implies f(x) \le f(y)$$

and is monotone decreasing if

$$x \le y \implies f(x) \ge f(y)$$

Proposition: Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex on $C \subset \mathbb{R}^n$.
Let $g : \mathbb{R} \to \mathbb{R}$ be convex and monotone increasing on range(f).
Then $g \circ f : \mathbb{R}^n \to \mathbb{R}$ is convex.

<u>Proof:</u> $g(f(\underline{x}))$ is convex ?

$$g(f(\lambda \underline{x} + (1-\lambda)\underline{y}))$$

$$f(\lambda \underline{x} + (1-\lambda)\underline{y}) \le \lambda f(\underline{x}) + (1-\lambda) f(\underline{y}) \quad \text{(f convex)}$$

$$g(\quad " \quad) \le g(\quad " \quad) \quad \text{(g monotone)}$$

$$g(f(\lambda \underline{x}) + (1-\lambda)\underline{y})) \le g(\lambda f(\underline{x}) + (1-\lambda) f(\underline{y}))$$

(g convex)

$$\le \lambda g(f(\underline{x})) + (1-\lambda) g(f(\underline{y})) \quad \blacksquare$$

$$\left.\begin{array}{l} f_1 = x^2 \\ f_2(x) = -x^{1/3} \end{array}\right\} \quad \text{range}(f_2) = (-\infty, 0]$$

on $(-\infty, 0] \Rightarrow f_1(x)$ is decreasing.

$\Longrightarrow$ can't apply this property.

Return to previous example:

<u>Ex</u>: Show $f(\underline{x}) = \sqrt{\|\underline{x}\|_2^2 + 1}$ is convex

Write $f(\underline{x}) = h(g(\underline{x}))$

$g(\underline{x}) = \|\underline{x}\|_2 \longrightarrow$ convex (it's a norm)

$h(t) = \sqrt{t^2+1}$

$$h'(t) = \frac{t}{\sqrt{t^2+1}} \geq 0 \quad \text{for } t \geq 0,$$

(ie. on range of g)

$$h''(t) = \frac{\sqrt{t^2+1} - t\frac{t}{\sqrt{t^2+1}}}{t^2+1}$$

$$= \frac{t^2+1-t^2}{(t^2+1)^{3/2}} = \frac{1}{(t^2+1)^{3/2}} > 0$$

$\Longrightarrow$ h is convex, monotone increasing.

$\Longrightarrow$ $f = h(g(\underline{x}))$ is convex.

Examples

Example (Beck, 7.23)

Show that $h(\boldsymbol{x}) = e^{\|\boldsymbol{x}\|^2}$ is convex.

$h(\underline{x}) = f_1(f_2(\underline{x}))$

$f_2(\underline{x}) = \|\underline{x}\| \longrightarrow$ convex

$f_1(t) = e^{t^2} \longrightarrow$ convex, increasing for $t \geq 0$

✓

Examples

Example (Beck, 7.24)

Show that $h(\boldsymbol{x}) = (\|\boldsymbol{x}\|^2 + 1)^2$ is convex.

$h(\underline{x}) = f_1(f_2(\underline{x}))$

$f_2(\underline{x}) = \|\underline{x}\|^2 \longrightarrow$ convex

$f_1(t) = (t+1)^2 \longrightarrow$ convex, increasing for $t \geq 0$

Thurs, Nov. 11

- Office hours Friday moved to 4:30-5:30pm (this week only)

- Final exam is on Wed Dec. 15 @ 8:00 am

- HW #5 due Tuesday.

Example: $f(\underline{x}) = \dfrac{\|\underline{A}\,\underline{x} + \underline{b}\|_2^2}{\underline{w}^T\underline{x} + c}$, $\underline{A}$, $\underline{b}$, $\underline{w}$, c are given/known.

("quadratic-over-linear")

consider related function: $g(\underline{x}, t) = \dfrac{\|\underline{x}\|_2^2}{t}$

$$= \sum_{i=1}^{n} \frac{x_i^2}{t}$$

Define $g_i(x_i, t) = \dfrac{x_i^2}{t}$.

From last week: $g_i(x_i, t)$ is convex.

$g(\underline{x}, t) = \sum_{i=1}^{n} g_i(x_i, t)$, g_i is convex

Since g is a conic combination of $\{g_i\}_{i=1}^{n}$,

then g is convex.

Finally: $f(\underline{x}) = g(\underbrace{\underline{\underline{A}}\,\underline{x} + \underline{b}}_{\text{affine in } \underline{x}}, \underbrace{c + \underline{w}^T \underline{x}}_{\text{affine in } \underline{x}})$

$\Rightarrow$ f is the composition of a convex function w/ an affine one

$\Rightarrow$ f is convex.

Pointwise maximum

Proposition: Suppose $f_i : C \to \mathbb{R}$, $C \subset \mathbb{R}^n$ convex.

Assume f_i is convex for $i = 1, \ldots, k$.

Then $f(x) = \max_{i=1..k} f_i(x)$ is convex.

Proof: First, note that

$$\max_{x \in \mathbb{R}} \left(g(x) + h(x) \right) \leq \max_{x \in \mathbb{R}} g(x) + \max_{x \in \mathbb{R}} h(x)$$

Let $\underline{x}, \underline{y} \in \cancel{\mathbb{R}^n}$ C, $\lambda \in (0,1)$

$$f(\lambda \underline{x} + (1-\lambda)\underline{y}) = \max_{i=1..k} f_i(\lambda \underline{x} + (1-\lambda)\underline{y})$$

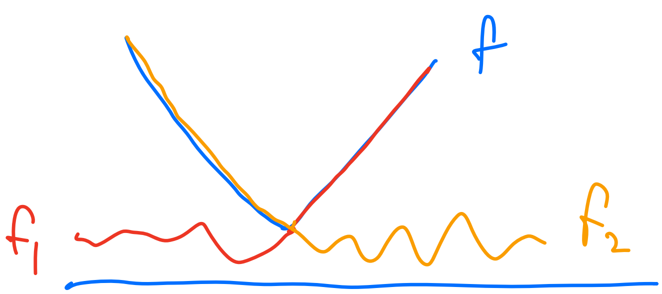
$$\leq \max_{i=1..k} \lambda f_i(\underline{x}) + (1-\lambda) f_i(\underline{y})$$

$$\leq \max_{i=1..k} \lambda f_i(\underline{x}) + \max_{i=1..k} ((1-\lambda) f_i(\underline{y}))$$

$$= \lambda \max_{i=1..k} f_i(\underline{x}) + ((1-\lambda) \max_{i=1..k} f_i(\underline{y}))$$

$$= \lambda f(\underline{x}) + (1-\lambda) f(\underline{y}). \quad \blacksquare$$

Converse not true:



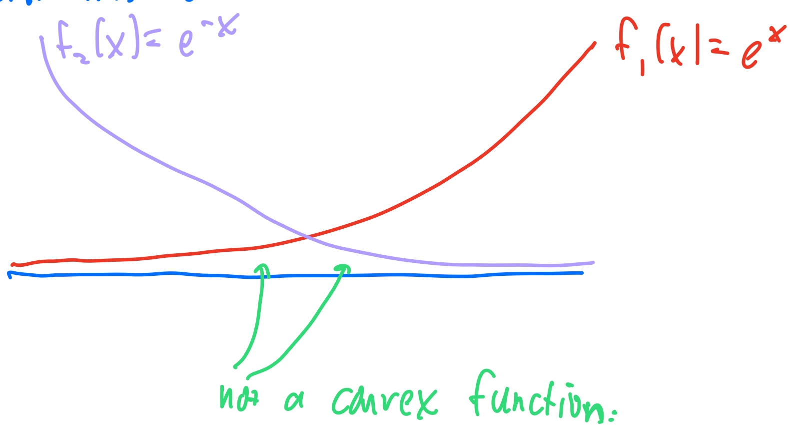
Ex. $f_1(\underline{x}) = \max_{i=1..n} x_i$ $\qquad$ $f_2(\underline{x}) = \|\underline{x}\|_\infty$

$$= \max_{i=1..n} |x_i|$$

f_1 is convex since x_i is convex ($\forall i$)

f_2 is convex since $|x_i|$ is convex

Minimums of convex functions need not be convex:



Level sets of convex functions

Definition: Given $f : \mathbb{R}^n \to \mathbb{R}$, then

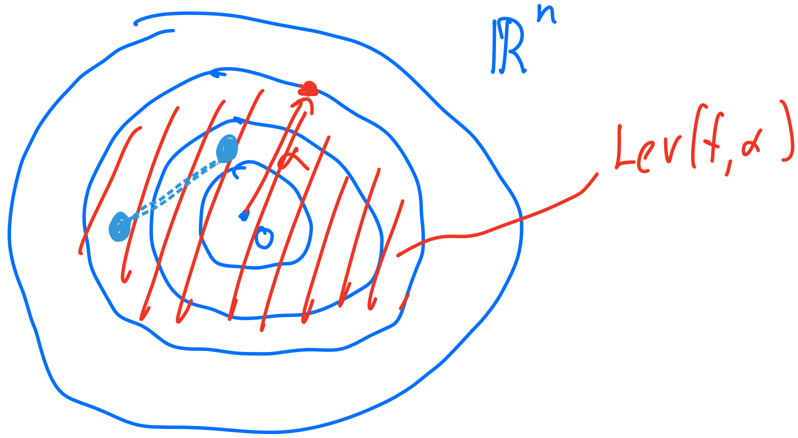
$$\mathrm{Lev}(f, \alpha) = f^{-1}((-\infty, \alpha]) = \{ \underline{x} \in \mathbb{R}^n \mid f(\underline{x}) \leq \alpha \}$$

($\alpha \in \mathbb{R}$ is arbitrary).

("Level set")

Proposition: If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then $\forall \alpha \in \mathbb{R}$, $\mathrm{Lev}(f, \alpha) = f^{-1}((-\infty, \alpha])$ is a convex set in $\mathbb{R}^n$.

$f(\underline{x}) = \|\underline{x}\|_2^2$



Proof: Suppose $\underline{x}, y \in Lev(f, \alpha)$. Let $\lambda \in (0, 1)$.

Then: $f(\underline{x}) \le \alpha$, $f(y) \le \alpha$

Is $\lambda \underline{x} + (1-\lambda) y \in Lev(f, \alpha)$?

$$f(\lambda \underline{x} + (1-\lambda) y) \le \lambda f(\underline{x}) + (1-\lambda) f(y)$$
$$\le \lambda \alpha + (1-\lambda)\alpha$$
$$= \alpha$$

$\Longrightarrow \lambda \underline{x} + (1-\lambda) y \in Lev(f, \alpha)$ ∎

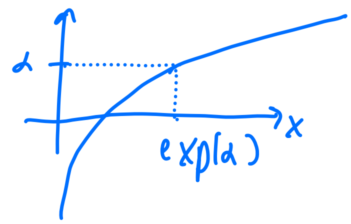
Functions with convex level sets

<u>Definition</u>: Let $f: \mathbb{R}^n \to \mathbb{R}$, f is called "quasi-convex" if $\mathrm{Lev}(f, \alpha)$ is convex $\forall\ \alpha \in \mathbb{R}$.

<u>Proposition</u>: f convex $\Longrightarrow$ f quasi-convex.

Converse is not true: f quasi-convex $\not\Longrightarrow$ f convex.

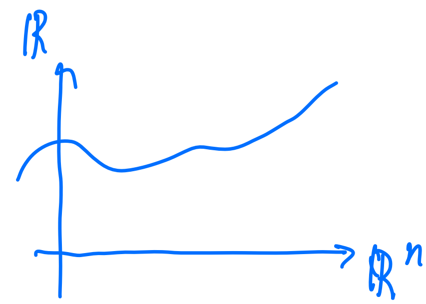
<u>Ex</u>: $f(x) = \log x$



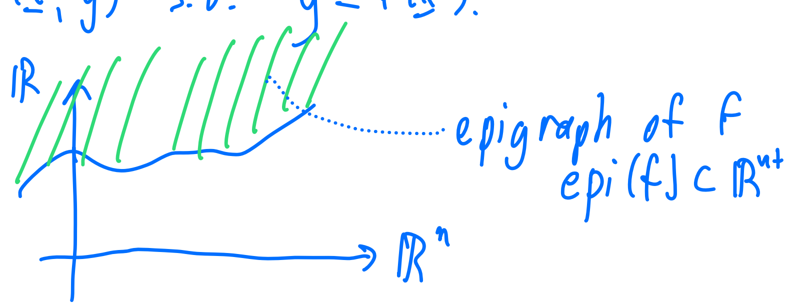
$Lev(f,\alpha) = (0, \exp(\alpha)) \to$ convex $\forall \alpha$.

$\Longrightarrow f$ quasi-convex. But f is not convex.

Let $f: \mathbb{R}^n \to \mathbb{R}$. The graph of f is the set of points $(\underline{x}, y) \in \mathbb{R}^{n+1}$ such that $y = f(\underline{x})$



The epigraph (or supergraph) of f is the set of points $(\underline{x}, y)$ s.t. $y \geq f(\underline{x})$.



Proposition : $f: \mathbb{R}^n \to \mathbb{R}$ is convex iff $epi(f)$ is convex in $\mathbb{R}^{n+1}$.

(✱: there are some pathological issues to deal with above.)

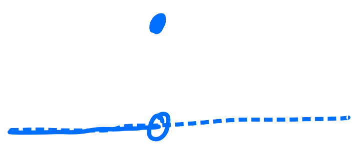
Convex functions and smoothness

Convex functions are not necessarily smooth.

Theorem: Let $f: C \to \mathbb{R}$, $C \subset \mathbb{R}^n$ is convex. Assume f is convex. Then for any $x \in \operatorname{int}(C)$, f is continuous at x.

(C°)

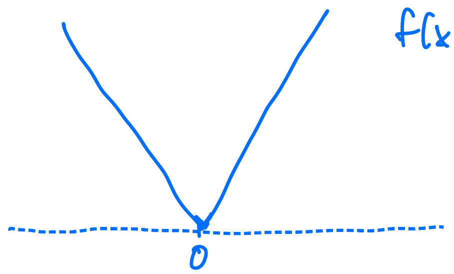
"Counterexample":



Theorem: Let $f: C \to \mathbb{R}$, $C \subset \mathbb{R}^n$ is convex. Assume f is convex. For any $x \in \operatorname{int}(C)$ and any $\underline{d} \in \mathbb{R}^n$, then

$\nabla f(\underline{x})^T \underline{d} = f'(\underline{x}; \underline{d})$ exists. (Directional derivatives exist.)

"Counterexample":



$f(x) = |x|$

$$\text{theorem guarantees this} \left\{ \begin{array}{l} \lim\limits_{h \downarrow 0} \dfrac{f(x+h) - f(x)}{h} \text{ exists} \\[2ex] \lim\limits_{h \uparrow 0} \dfrac{f(x+h) - f(x)}{h} \text{ exists} \end{array} \right.$$

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \text{ does not exist.}$$