# Lecture 4

## Last time

- ▷ Convex functions
- ▷ Continuity
- ▷ Gradients

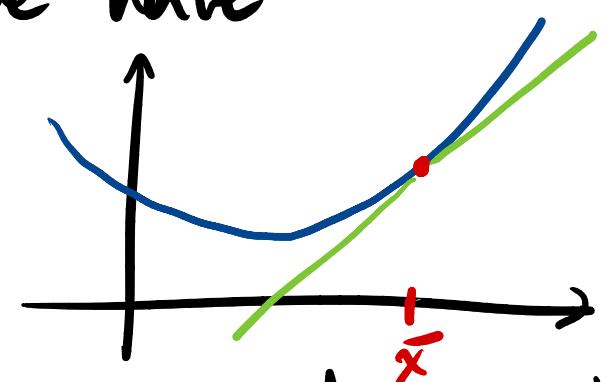
## Today

- ▷ Subgradients
- ▷ Normals
- ▷ Optimality conditions with convex sets

## Subgradients

How do we generalize gradients for nonsmooth convex functions?

Recall that for smooth convex functions we have




$$f(x) \geq f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x} \rangle.$$

Inspired by this property:

**Def:** Let $f : E \to \overline{\mathbb{R}}$ be a convex function and let $\bar{x} \in \operatorname{dom} f$. The set of subgradients (or subdifferential) of $f$ at $\bar{x}$ is

$$\partial f(\bar{x}) = \{ g \in E \mid f(\bar{x}) + \langle g, y - \bar{x} \rangle \leq f(y) \ \forall y \}.$$

$\dashv$

This concept generalizes gradients.

**Proposition:** Let $f : E \to \bar{\mathbb{R}}$ be a convex function. If $f$ is differentiable at $\bar{x}$, then, $\partial f(\bar{x}) = \{ \nabla f(\bar{x}) \}$. $\dashv$

Moreover, they give simple optimality conds.

**Lemma (unconstrained optimality condition)** Let $f : E \to \bar{\mathbb{R}}$ be convex. Then, $x^*$ is a minimizer if, and only if, $0 \in \partial f(x^*)$. $\dashv$

Q. But do subgradients exist?

**Theorem (Existance of subgradients).** Let $f : E \to \bar{\mathbb{R}}$ be a convex function and $\bar{x} \in \operatorname{int} \operatorname{dom} f$. Then, $\partial f(\bar{x})$ is not empty.

Previous lecture.

**Proof:** By Corollary (4), since

$f : \operatorname{dom} f \to \mathbb{R}$ is convex, we have $f$ is Lipschitz at $\bar{x}$. Consider the sequence

$$(\bar{x}, f(\bar{x}) - 1/n) \notin \operatorname{epi} f,$$

therefore $(\bar{x}, f(\bar{x})) \in \operatorname{bd} \operatorname{epi} f$.

Using that $f$ is locally Lipschitz we can show $(\bar{x}, f(\bar{x}) + 1) \in \operatorname{int} \operatorname{epi} f$ (Why?), thus $\operatorname{int} \operatorname{epi} f \neq \emptyset$.

Hence, Hahn-Banach ensures the existence of $a = (h, \gamma) \in E^* \times \mathbb{R}$ s.t.

$$\langle a, (\bar{x}, f(\bar{x})) \rangle \leq \langle a, (x, t) \rangle \quad \forall (x, t) \in \operatorname{epi} f.$$

Moreover, the inequality is strict for $(x, t) \in \operatorname{int} \operatorname{epi} f$ (Why?). Since $(\bar{x}, f(\bar{x}) + 1)$ is in the interior of $\operatorname{epi} f$, we conclude $0 < \gamma$. Then we can rescale $\bar{a} = \frac{1}{\gamma} a$ to obtain that

$$\langle \bar{h}, \bar{x} \rangle + f(\bar{x}) \leq \langle \bar{h}, x \rangle + f(x) \quad \forall x \in E$$

$$\Updownarrow$$

$$f(\bar{x}) + \langle (-\bar{h}), x - \bar{x} \rangle \leq f(x) \quad \forall x \in E.$$

Thus, $(-\bar{h}) \in \partial f(\bar{x})$. $\square$

## First order optimality conditions

We come back to one of our problems of interest:

$$\min \; f(x) \quad \text{s.t.} \; x \in C$$

$\leftarrow$ convex and closed.

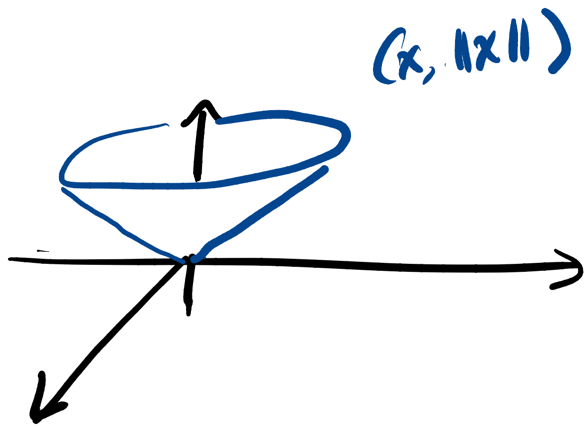
A critical quantity for our conditions will be the directional derivative.

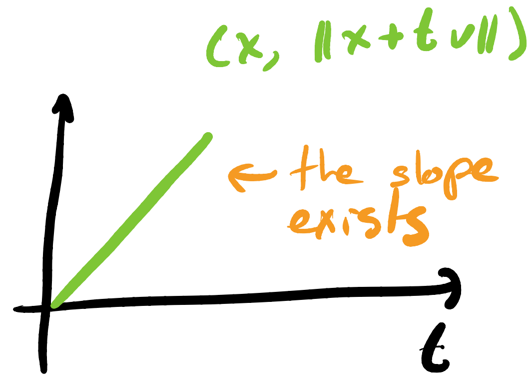
**Def:** Given a function $f: \mathbb{E} \to \bar{\mathbb{R}}$ a point $\bar{x} \in \operatorname{dom} f$ and a direction $v \in \mathbb{E}^*$. We say that $f$ is directionally differentiable at $\bar{x}$ in the direction $v$ if the following limit exists.

$$f'(x; v) = \lim_{t \downarrow 0} \frac{f(\bar{x}+tv) - f(\bar{x})}{t}.$$

## Example

- The norm $\|\cdot\|$ is not differentiable at zero. But it is directionally differentiable for all $v \in \mathbb{E}^*$.




$\dashv$

**Lemma:** The following two hold.

1) If $f : E \to \mathbb{R}$ is differentiable, then
$$f'(x; v) = \langle \nabla f(x), v \rangle \qquad \forall x \in E,\ v \in E^*.$$

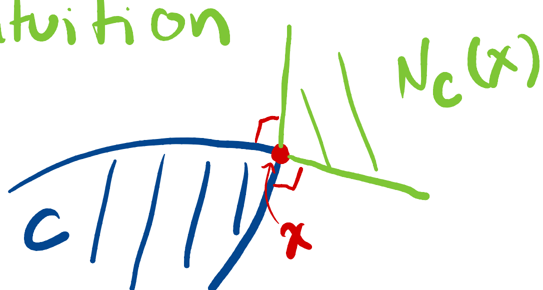
2) If $f : E \to \overline{\mathbb{R}}$ is convex, then
$$f'(x; v) = \sup_{g \in \partial f(x)} \langle g, v \rangle \qquad \forall x \in \operatorname{int} \operatorname{dom} f,\ v \in E^*.$$

**Proof:** Exercise. $\square$

We need one extra ingredient.

**Def:** Given a closed, convex set $C$. The **normal cone** of $C$ at $\bar{x} \in C$ is given by:
$$N_C(\bar{x}) = \{ g \in E \mid \langle g, x - \bar{x} \rangle \leq 0 \quad \forall x \in C \}.$$
If $\bar{x} \in C$ we let $N_C(\bar{x}) = \emptyset$. $\dashv$

Intuition



In HW 1 you'll prove that $N_C(\bar{x})$ is a closed convex cone (and some extra properties).

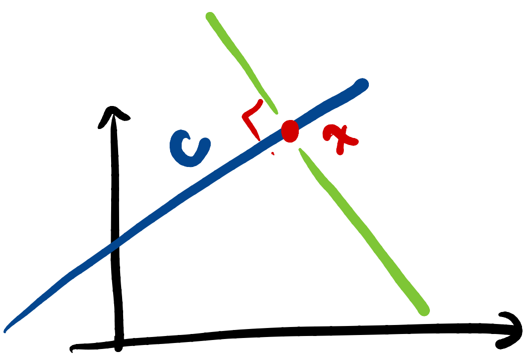
## Examples

- Subspace $C = \{x \mid Ax = b\}$.

$b \in F$

Linear map $A : E \to F$

another Euclidean space




Then for $\bar{x} \in C$ we have

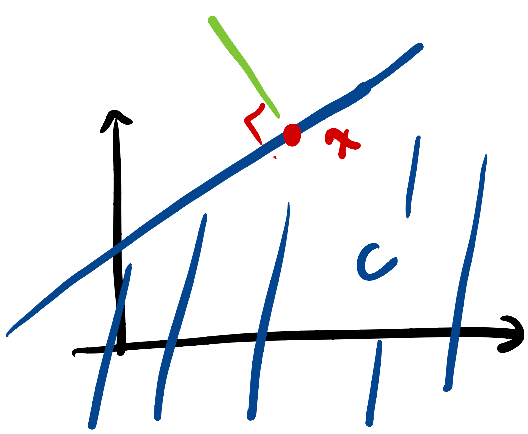
$$N_C(\bar{x}) = \{A^* y \mid y \in F\}.$$

Adjoint of $A$

i.e., $\langle Ax, y\rangle = \langle x, A^* y\rangle \quad \forall x, y.$

- Half space $C = \{x \mid \langle a, x\rangle \le \beta\}$

$a \in E^*$ $\qquad$ $\beta \in \mathbb{R}.$

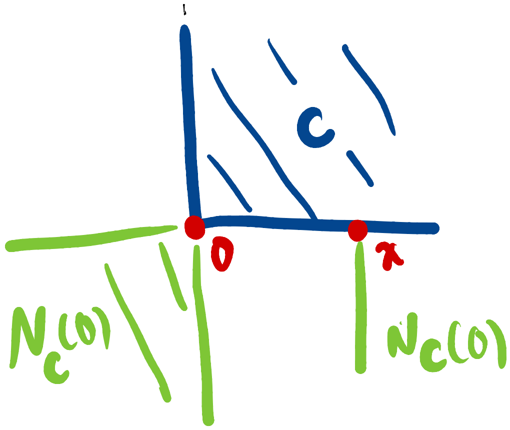



Then, for $\bar{x} \in C$,

$$N_C(\bar{x}) = \begin{cases} \{0\} & \text{if } \langle a, \bar{x}\rangle < \beta, \\ \{\lambda a \mid \lambda \ge 0\} & \text{otherwise.} \end{cases}$$

- Nonnegative orthant

$$C = \mathbb{R}^d_+ = \{x \in \mathbb{R}^d \mid x_i \ge 0 \quad \forall i \in [d]\}.$$



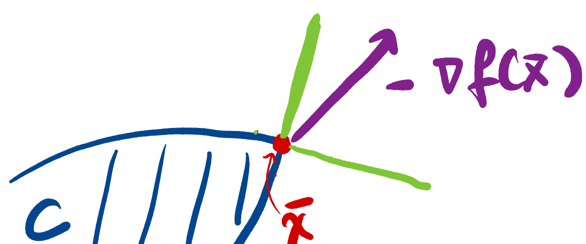

Then, for $\bar{x} \in C$

$$N_C(\bar{x}) = \left\{ g \in \mathbb{R}^d \;\middle|\; \begin{array}{ll} g_i = 0 & \text{if } \bar{x}_i > 0 \\ g_i \leq 0 & \text{if } \bar{x}_i = 0 \end{array} \right\}.$$

⊣

**Proposition** (Necessary condition): Suppose $C \subseteq E$ closed and convex and $\bar{x} \in$ is a local minimizer of $f$ over $C$. Then, if $f'(\bar{x}; x - \bar{x})$ exists for some $x \in C$, it has to be nonnegative. In particular, if $f$ is differentiable a $\bar{x}$, then $-\nabla f(\bar{x}) \in N_Q(\bar{x})$. ⊣

Intuition




**Proof**: Seeking contradiction suppose $\exists\, x \in C$ s.t. $f'(\bar{x}, x - \bar{x}) < 0$. Then, there is a $\delta > 0$ sufficiently small s.t. $\forall\, t \in (0, \delta)$ we have

$$\frac{f(\bar{x}+t(x-\bar{x}))-f(\bar{x})}{t} \leq 0$$

$$\Rightarrow \quad f(\bar{x}+t(x-\bar{x})) \leq f(\bar{x}).$$

Since $C$ is convex, we have $\bar{x}+t(x-\bar{x}) \in C$.
Therefore $\bar{x}$ is not a local minimizer. $\lightning$
When $f$ is differentiable

$$0 \leq f'(\bar{x}; x-\bar{x}) = \langle \nabla f(\bar{x}), x-\bar{x} \rangle$$

$$\Leftrightarrow \quad -\nabla f(\bar{x}) \in N_C(\bar{x}).$$

This completes the proof. $\square$

The converse is not true, in general. But it holds if we assume $f$ is also convex.

Proposition (sufficient condition):
Suppose $C$ and $f$ are closed and convex. Suppose that for $\bar{x} \in C$ we have

$$f'(\bar{x}; x-\bar{x}) \geq 0 \quad \forall x \in C.$$

Then, $\bar{x}$ is a minimizer of $f$ over $C$. In particular if $f$ is differentiable at $\bar{x}$,

$$-\nabla f(\bar{x}) \in N_C(\bar{x}) \quad \Rightarrow \quad \bar{x} \in \operatorname*{argmin}_{x \in C} f(x). \quad \dashv$$

**Proof:** Recall a claim from the previous lecture:

Claim($\star$): Suppose that $g : \mathbb{R}^+ \to \mathbb{R}$ is convex with $g(0)=0$. Then, $t \mapsto \frac{g(t)}{t}$ is nondecreasing. $\dashv$

For any $x \in C$, the function

$g_x(t) = f(\bar{x} + t(x - \bar{x})) - f(\bar{x})$

satisfies that $g_x(t)/t$ is nondecreasing. Thus, by assumption, for any sufficiently small $t$ we have

$$0 \leq \frac{g_x(t)}{t} \underset{\text{Claim } (\star)}{\leq} \frac{g_x(1)}{1} = f(x) - f(\bar{x})$$

Thus, $f(\bar{x}) \leq f(x)$ for all $x \in C$. $\square$