Lecture 3 Sep/03/2024

Recap

- Optimality conditions
  - ↳ First order neccessary cond.
  - ↳ First order sufficient cond. for convex functions.

Agenda

...
  - ↳ second order nessary cond.
  - ↳ Second order sufficient cond.
- Basic convex analysis.
  - → Convex sets
  - → smooth convex functions
  - → subdifferentials.

Optimality conditions

Theorem (1st-order sufficient condition)

Assume that $f: \mathbb{R}^d \to \mathbb{R}$ is a smooth convex function.
Then, x^* satisfies $\nabla f(x^*) = 0$ iff x^* is a global minimizer.

Proof: "$\Leftarrow$" Done. ✓

"$\Rightarrow$" Assume $\nabla f(\bar{x}^*) = 0$. Then,

Let $\bar{y} \in \mathbb{R}^d \setminus \{x^*\}$.

Define $\phi(t) = f(\bar{x}^* + t(\bar{y} - \bar{x}^*))$.

By chain rule

$$\phi'(0) = (\bar{y} - x^*) \nabla f(x^*) = 0$$

For any $t \in [0, 1]$ we have

$$\frac{f(\bar{x}^* + t(\bar{y} - \bar{x}^*)) - f(x^*)}{t \|\bar{y} - \bar{x}^*\|}$$

(convexity) $$\leq \frac{(1-t) f(\bar{x}^*) + t f(\bar{y}) - f(x^*)}{t \|\bar{y} - \bar{x}^*\|}$$

$$= \frac{\cancel{t}(f(\bar{y}) - f(\bar{x}^*))}{\cancel{t} \|\bar{y} - \bar{x}^*\|}$$

Taking limits on both sides

$$0 = \phi'(0) \leq \frac{f(\bar{y}) - f(\bar{x}^*)}{\|\bar{y} - \bar{x}^*\|}. \quad \square$$

Theorem (2nd-order nessary cond)

Suppose $f:\mathbb{R}^d \to \mathbb{R}$ twice diff (C^2).

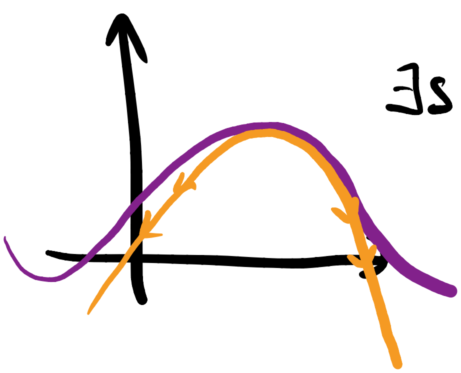
If $\bar{x}^*$ is a local min

$\Rightarrow \nabla f(\bar{x}^*) = 0$ and $s^T \nabla^2 f(\bar{x}^*) s \geq 0 \quad \forall s \in \mathbb{R}^d$

$\nabla f^2(\bar{x}^*)$ is positive semidefinite

$\nabla^2 f(\bar{x}^*) \succeq 0$.

Intuition



If $\nabla f(\bar{x}^*) = 0$, but

$\exists s$ s $\nabla^2 f(x^*) s < 0$

$\Rightarrow$ Exists direction downhill (2nd-order)

$\Rightarrow$ Better point.

Proof: Suppose seeking contradiction

$\nabla f(x^*) = 0$ and $\exists \bar{s} : \bar{s}^T \nabla f(\bar{x}^*) \bar{s} < 0$.

Define: $\phi(t) = f(\bar{x}^* + t\bar{s})$.

norm 1.

Then by def

$$0 > \frac{1}{2}\phi''(0) = \lim_{t\to 0} \frac{\phi(t)-\phi(0)}{t^2}$$

For small enough $t>0$

$$0 > \frac{1}{4}\phi''(0) \geq \frac{\phi(t)-\phi(0)}{t^2}$$

$$\Rightarrow \quad f(\bar{x}^*) > f(\bar{x}^* + t\bar{s}).$$

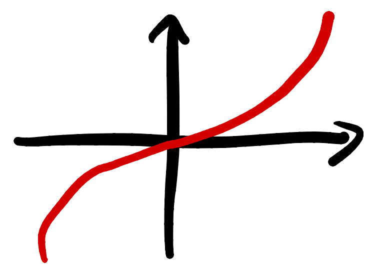
↯ $\square$

Is this sufficient?

No! Same example as before

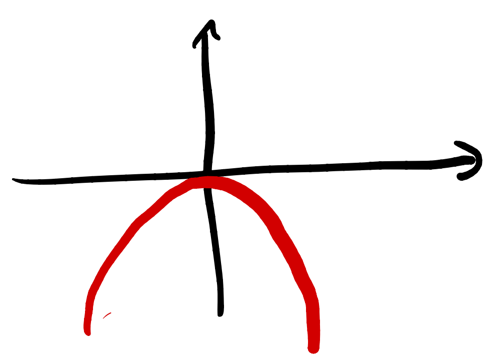
$f(x) = x^3$

$f''(0) = 0$



$f(x) = -x^4$

$f''(0) = 0$



Ex: Come up with a 2D example

where $\nabla^2 f(0) \neq 0$. $\dagger$

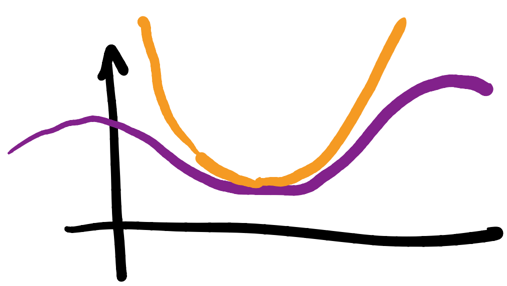
Theorem: 2^{nd}-order sufficient cond.

Suppose $f: \mathbb{R}^d \to \mathbb{R}$ is twice diff. If $\bar{x}^* \in \mathbb{R}^d$ satisfies that

$\nabla f(\bar{x}^*) = 0$
$\nabla^2 f(\bar{x}^*) \succ 0$ $\Rightarrow$ x^* is a strict local minimum

$s^T \nabla^2 f(\bar{x}^*) s > 0 \quad \forall s \in \mathbb{R}^d$.

Intuition:



The function curves upwards in every direction $\Rightarrow$ x^* is strict local minimum.

Proof: Suppose $\bar{x}^*$ satistfies the assumptions.

Let $\bar{u} \in \mathbb{R}$, with $\|\bar{u}\| = 1$

Let $\varphi(s) = f(\bar{x}^* + s\bar{u})$.

By the Fundamental Theorem of calculus

$$\varphi(s) = \varphi(s) + \int_0^s \varphi'(\alpha)\, d\alpha$$

Applying it again on $\varphi'(t)$

$$\varphi(s) = \varphi(0) + \varphi'(0) + \int_0^s \int_0^\alpha \varphi''(\beta)\, d\beta\, d\alpha$$

HW 1

Since $\nabla^2 f(x^*)$ is continuous and $\lambda := \lambda_{min}(\nabla f(x^*)) > 0$, then for all points y close to x^*

$$\lambda_{min}(\nabla f(y)) \geq \frac{\lambda}{2}.$$

Then, for small enough s

$$\varphi(s) = \varphi(0) + \varphi'(0) + \int_0^s \int_0^\alpha \bar{u}^T \nabla^2 f(\bar{x}^* + \beta \bar{u})\, u \, d\beta\, d\alpha$$

$$\geq \psi(0) + \frac{\lambda}{2}\int_0^s\int_0^\alpha 1 \, d\beta \, d\alpha$$

$$= \psi(0) + \frac{\lambda s^2}{4}$$

$$> \psi(0)$$

$$\Rightarrow \quad f(\bar{x}^*) = \psi(0) < \psi(s) = f(\bar{x}^* + s\bar{u}).$$

any point in a near by radius.

$\square$

Basics of convexity

We already saw convex functions

<u>Def</u>: $f(t\bar{x} + (1-t)\bar{y}) \leq t f(\bar{x}) + (1-t) f(\bar{y})$ $\forall \bar{x}, \bar{y}, \; t \in [0,1].$

There is also a natural notion of convexity for sets

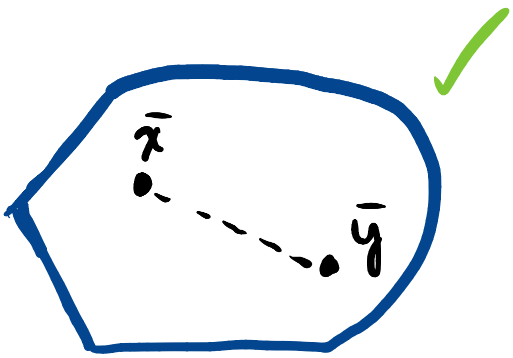
<u>Def</u> A set $C \subseteq \mathbb{R}^d$ is **convex** if

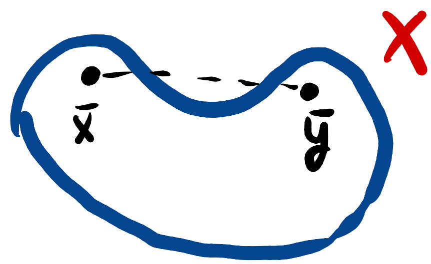
for all $\bar{x}, \bar{y} \in C$ and $t \in [0,1]$

$t\bar{x} + (1-t)\bar{y} \in C$.

Intuition

Segments never go out!



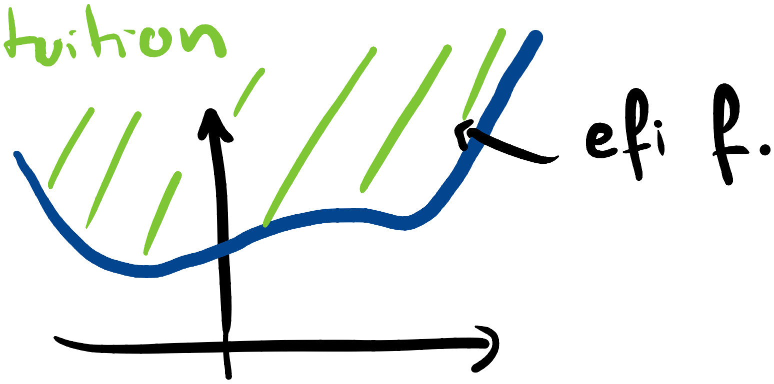


There is a deep connection between convex sets and convex functions.

Def: Given a convex function $f: \mathbb{R}^d \to \mathbb{R}$, its **epigraph** is given by

$$\text{epi } f := \{(x,t) \mid f(x) \le t\}$$

Intuition



Proposition: A function is convex iff its epigraph is convex.

Proof: HWL. $\dashv$

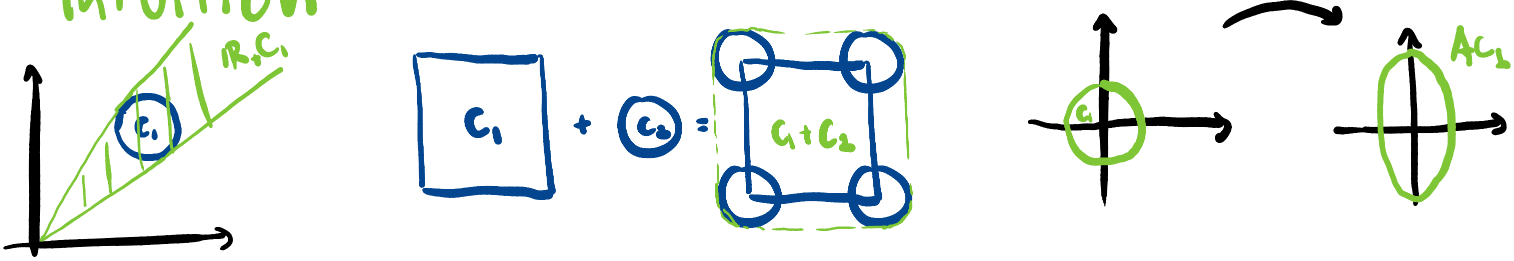
The relationship between convex functions go deeper than this. If you are interested consider taking "Intro to Convexity" with Amitabh Basu.

$C_3 \in \mathbb{R}^n$

Lemma: Assume that $C_1, C_2 \in \mathbb{R}^d$ convex sets. Then, the following are convex

1. (Scaling) $\mathbb{R}_+ C_1 = \{\lambda \bar{x} \mid \lambda \geq 0 \text{ and } x \in C_1\}$
2. (Sums) $C_1 + C_2 = \{\bar{x}_1 + \bar{x}_2 \mid x_1 \in C_1, x_2 \in C_2\}$
3. (Intersections) $C_1 \cap C_2$.
4. (Linear images and preimages)
   Let $A : \mathbb{R}^d \to \mathbb{R}^n$ is linear,
   $A C_1$ and $A^{-1} C_3$ are convex. $\dashv$

Intuition



Proof: Exercise $\square$

Equivalence of operations

Function	Epigraph
$\lambda f(x/\lambda)$	$\lambda \operatorname{epi} f$
$\max_i f_i$	$\bigcap_i \operatorname{epi} f_i$
$f(Ax)$	$[A \times I]^{-1} \operatorname{epi} f$