Week 3

Nicolo Michelusi

I. Convex optimization problems

- Standard form:

$$\min \ f_0(x) \rightarrow \text{objective function}$$
$$\text{s.t. } f_i(x) \leq 0 \ \forall i \rightarrow \text{inequality constraints}$$
$$h_j(x) = 0 \ \forall j \rightarrow \text{equality constraints}$$

$f_i(x)$ are convex ; $h_i(x) = a_i^T x + b_i$

- Domain of the problem:

$$D = \text{dom}(f_0) \cap \bigcap_i \text{dom} f_i \cap \bigcap_j \text{dom}(h_j)$$

- Feasible point/set

$x \in D$ and x satisfies constraints

i.e. : x feasible if $x \in D$ and $f_i(x) \leq 0 \ \forall i > 0$

$h_j(x) = 0 \ \forall j$

- Optimal value

$$p^* = \inf \{ f_0(x) \mid f_i(x) \le 0 \ \forall i \ ; \ h_j(x) = 0 \ \forall j \}$$

$p^* = +\infty$ if no point is feasible

$p^* = -\infty$ if $\exists$ x_k feasible s.t. $\lim_k f_0(x_k) = -\infty$

- Optimal point (or set of)

If p^* is attained at a feasible point x^*, i.e. $f_0(x^*) = p^*$

$\Rightarrow$ x^* is optimal

- More general form:

Let $C = \{ x : f_i(x) \le 0 \ \forall i \ge 0 \ ; \ h_j(x) = 0, \forall j \}$ (convex)

$$= \bigcap_{i \ge 0} \{ x : f_i(x) \le 0 \} \cap \bigcap_j \{ x : h_j(x) = 0 \}$$

$\Rightarrow$ problem equivalent to $\min_{x \in C} f_0(x)$

Q: why must h_i be affine?

$\{ x : h_1(x) = 0 \}$ must be convex

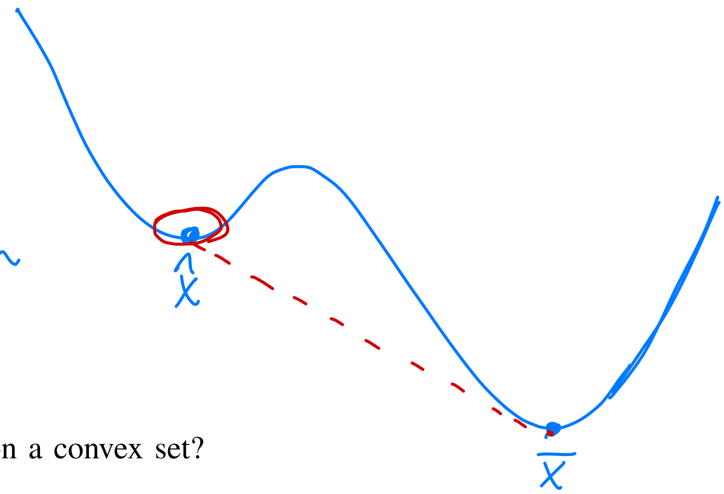
$\Rightarrow$ Let x_1, x_2, θ $\Rightarrow$ $h(x_1) = 0, h(x_2) = 0$

$$h(\theta x_1 + \bar{\theta} x_2) = 0$$

$\Rightarrow$ $h(\theta x_1 + \bar{\theta} x_2) = \theta h(x_1) + \bar{\theta} h(x_2)$ $\Rightarrow$ must be affine

II. Basic properties of convex problems

- Local/global optimum

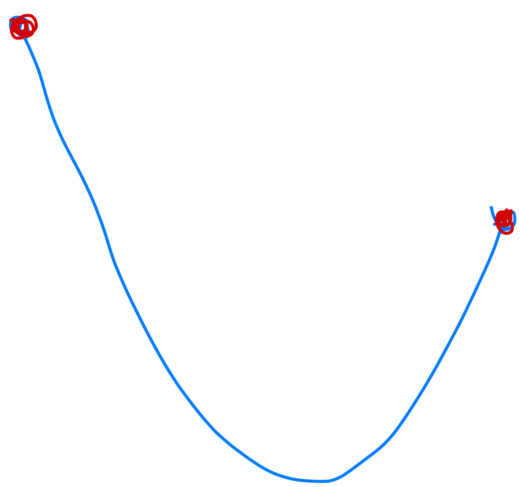


Local optimum $\hat{x} \in C$

$\exists r>0: f_0(\hat{x}) \le f_0(x) \quad \forall x \in C: \|x-\hat{x}\| \le r$

$\Longrightarrow$ in convex problem, a local optimum is also <u>global</u>

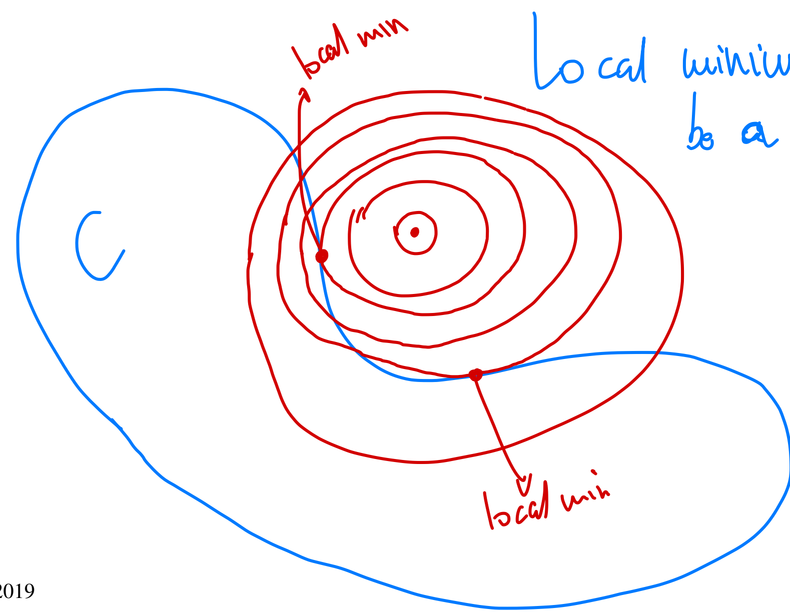
- what if we try to maximize convex function on a convex set?



optimal point is on the boundary of feasible set!

- what if we try to minimize convex function on a non-convex set? (local minimum at the boundary might not be global minimum)

Local minimum might not be a global one



III. Standard form

- We will use the standard from, where f_i are convex and h_i are linear. Many problems, even if they don't appear convex or in standard form at first glance, can be converted into convex problems in standard form with some tricks.
- Change of variables

$$\min f_0(x) \quad \text{s.t. } f_i(x) \le 0 \ \forall i = 1, \dots, p, \quad h_j(x) = 0 \ \forall j = 1, \dots, m \quad \xRightarrow{x = g(z)} \quad \min f_0(g(z)) \quad \text{s.t. } f_i(g(z)) \le 0 \ \forall i = 1, \dots, p, \quad h_j(g(z)) = 0 \ \forall j = 1, \dots, m$$

- Example: $\min_{x,y>0,z>0} x^2 + y$ s.t. $\ln(y+z) \le x, y^2 + z^2 \le 1$ (use $y = e^t, z = e^w$)

$$\min_{x,y>0,z>0} x^2 + y \quad \text{s.t. } \ln(y+z) \le x \ (\text{non convex}), \quad y^2 + z^2 \le 1 \quad \Longrightarrow \quad \min_{x,t,w} x^2 + e^t \quad \text{s.t. } \ln(e^t + e^w) - x \le 0 \ (\text{convex}), \quad e^{2t} + e^{2w} \le 1$$

- Example: $\min_{x,y} \sqrt{x} - \sqrt{y}$ s.t. $x + y \le 1$ (use $\sqrt{x} = z, \sqrt{y} = w$) (~~use $y = e^t, z = e^w$~~)

$$\min \sqrt{x} - \sqrt{y} \ (\text{non convex}) \quad \text{s.t. } x + y \le 1 \quad \Longrightarrow \quad \min z - w \quad \text{s.t. } z^2 + w^2 \le 1, \quad z \ge 0, \quad w \ge 0$$

- Transformation of functions
- Example: $\frac{x_1}{1+x_2^2} \cancel{\le} \ge 1$ (non convex)

$$\Longrightarrow \frac{x_1}{1+x_2^2} \ge 1 \iff -x_1 + 1 + x_2^2 \le 0 \ (\text{convex}) \Longrightarrow \text{feasible set is the same}$$

- Example: $\min \sqrt{x_1 + x_2}$ (concave)

$\min \sqrt{x_1+x_2} = \sqrt{\min x_1+x_2}$ since $\sqrt{\cdot}$ is monotonic

$\min x_1+x_2$ is convex

- Converting equality constraints into inequality constraints (if you know additional structure of the problem)
- Example: $\min x_1 + x_2$ s.t. $x_1^2 + x_2^2 = 1$ non convex

however equivalent to:

$\min x_1+x_2$

s.t. $x_1^2+x_2^2 \leq 1$ (convex)

since objective function is monotonic in x_1, x_2 and optimum is achieved at the boundary of feasible set

- Implicit constraints can be made explicit
- Example: $\min f(x)$ with $f(x) = x^T x$ for $Ax = b$ and $f(x) = +\infty$ otherwise

$\min \begin{cases} x^Tx & Ax=b \\ +\infty & \text{otherwise} \end{cases} \Rightarrow \min x^Tx$ s.t. $Ax=b$ (convex)

- Simplify objective/constraint functions by introducing additional constraints
- Example: $\min f_0(A_0x + b_0)$ s.t. $f_i(A_ix + b_i) \leq 0$

$\min f_0(A_0x+b)$ s.t. $f_i(A_ix+b_i) \leq 0 \Rightarrow \min f_0(y_0)$ s.t. $f_i(y_i) \leq 0$, $y_0 = A_0x+b_0$, $y_i = A_ix+b_i \ \forall i$

- Sometimes we can do so with non-linear mappings, e.g. $\min f_0(g(x))$

$$\min f_0(g(x)) \equiv \min_{x,y} f_0(y) \quad \text{s.t. } g(x)=y$$

If additionally f_0 increasing

$$\Rightarrow \min_{x,y} f_0(y) \quad \text{s.t. } g(x) \le y$$

- Conversely, equality constraints can be absorbed into the function
- Example: $\min f_0(x)$ s.t. $f_i(x) \le 0, Ax = b$

$A \in \mathbb{R}^{m \times n}$, $m \le n$, A full rank $\Rightarrow Ax=b \Leftrightarrow \exists z \in \mathbb{R}^{n-m}$ s.t. $x = x_0 + F \cdot z$,

where F spans the null space of A, $F \in \mathbb{R}^{n \times n-m}$, $AF=0$ and x_0 is one solution of $Ax_0 = b$

$$\Rightarrow \min f_0(x) \quad \text{s.t. } f_i(x) \le 0 \ \forall i, \ Ax=b \quad \equiv \quad \min_{z \in \mathbb{R}^m} f_0(x_0 + Fz) \quad \text{s.t. } f_i(x_0 + Fz) \ \forall i$$

- Epigraph form

$$\min f_0(x) \quad \text{s.t. } f_i(x) \le 0 \ \forall i, \ h_j(x) = 0 \ \forall j \quad \equiv \quad \min t \quad \text{s.t. } f_0(x) - t \le 0, \ f_i(x) \le 0 \ \forall i, \ h_j(x) = 0 \ \forall j$$

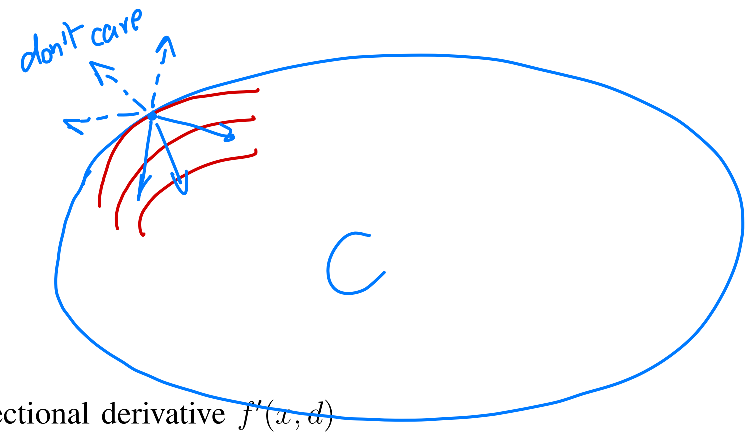
- Slack variables to convert inequality constraints into equality ones

$$\min f_0(x) \quad \text{s.t. } f_i(x) \le 0 \quad \equiv \quad \min f_0(x) \quad \text{s.t. } f_i(x) - s_i = 0, \ s_i \le 0 \quad \forall i$$

- All of these techniques may help to transform an otherwise non-convex problem into a convex one.

IV. Conditions for optimality

- When is a point x optimal? Roughly speaking, when the function is non-decreasing in any direction pointing towards the interior of the feasible set:



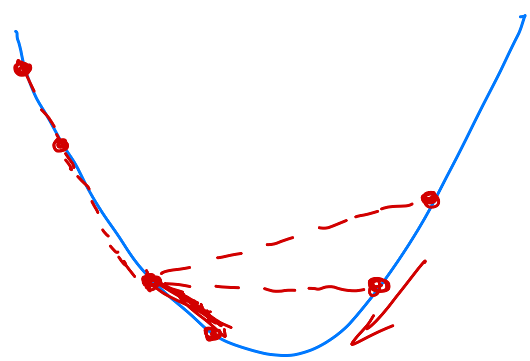
- Directional derivative $f'(x, d)$

$$f'(x,d) = \lim_{t \to 0^+} \frac{f(x+td) - f(x)}{t}$$

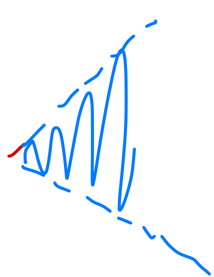
- If $f'(x, d) = a^T d$ for all directions d, then the function is said to be differentiable at x, with gradient $\nabla f(x) = a$
- Existence of directional derivatives for convex function (possibly, $= \pm\infty$):

If $f: C \to \mathbb{R}$ is convex

$\Longrightarrow$ $f'(\bar{x}, x - \bar{x})$ exists $\forall\ x, \bar{x} \in C$



- May not exist for arbitrary functions, e.g. $x \sin(1/x)$ (does not exist in 0)

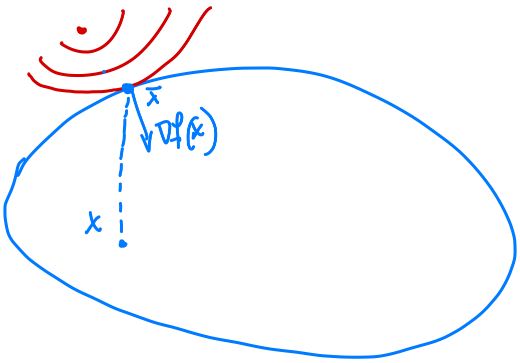


$f(x) = x \sin\left(\frac{1}{x}\right) \rightarrow \lim_{x \to 0} f(x) = 0$

$f'(x) = \sin\left(\frac{1}{x}\right) - \frac{1}{x}\cos\left(\frac{1}{x}\right) \rightarrow \nexists$ for $x \to 0$

V. Necessary conditions for optimality

- Necessary condition for optimality: let $\bar{x}$ be a local minimizer of $f(x)$ in $\mathcal{C}$ convex; then, $f'(\bar{x}, x-\bar{x}) \geq 0$, $\forall x \in \mathcal{C}$
  (holds for any local minimum)
- What if the problem is convex? (local min is also global)
- Necessary condition for f differentiable



$$f'(\bar{x}; x-\bar{x}) = \nabla f(\bar{x})^T (x-\bar{x}) \geq 0 \quad \forall x \in C$$

<u>Proof</u>: $\exists\, r>0 : f(\bar{x}) \leq f(z) \ \forall z \in C$ s.t. $\|z-\bar{x}\| \leq r$

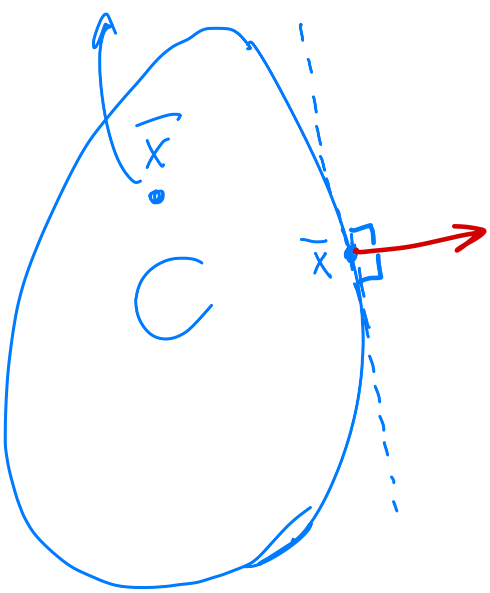
Let $x \in C \Rightarrow \exists\, \theta \in (0,1)$ s.t. $z = \theta x + \bar{\theta}\bar{x} \in C$ and $\|z-\bar{x}\| \leq r$ (true for $\theta \to 0$)

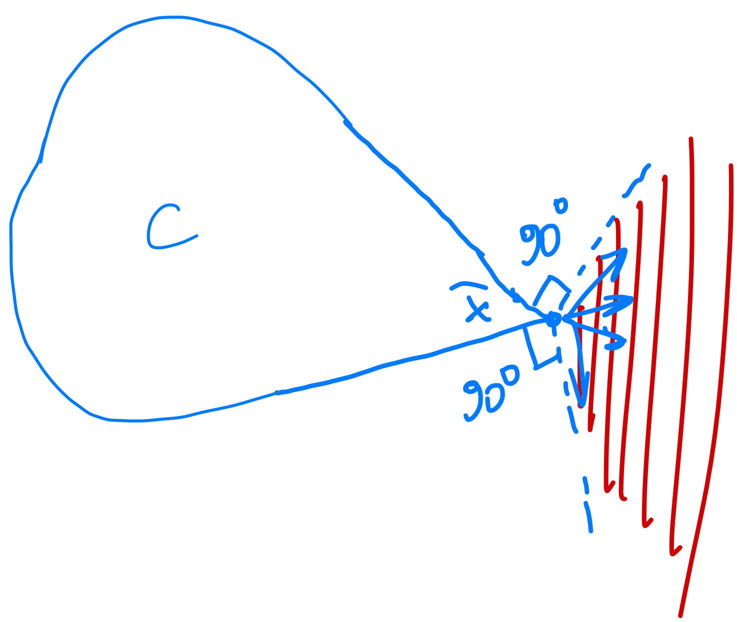
$$\Rightarrow f(z) = f(\theta x + \bar{\theta}\bar{x}) \geq f(\bar{x}) \Rightarrow \frac{f(\bar{x}+\theta(x-\bar{x})) - f(\bar{x})}{\theta} \xrightarrow{\theta \to 0} \nabla f(\bar{x})^T (x-\bar{x}) \geq 0$$

- To further understand these conditions, define the normal cone $N_C(\bar{x})$ to a convex set C at point $\bar{x}$:

$$N_C(\bar{x}) = \{d : d^T(x-\bar{x}) \leq 0, \forall x \in C\}$$

$N_C(\bar{x}) = \{0\}$





- With this definition, necessary condition of optimality for f differentiable becomes

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

$\nabla f(\bar{x})^T(x-\bar{x}) \geq 0 \quad \forall x \in C; \Rightarrow -\nabla f(\bar{x})^T(x-\bar{x}) \leq 0 \quad \forall x \in C$

$\Rightarrow -\nabla f(\bar{x}) \in N_c(\bar{x})$

In particular, if $\bar{x} \in \text{int}(C) \Rightarrow N_c(\bar{x}) = \{0\} \Rightarrow \nabla f(\bar{x}) = 0$

VI. Sufficient conditions for optimality (convex functions only)

- First order sufficient condition: let f convex and C convex; let $\bar{x} \in C$. Then, if $f'(\bar{x}; x-\bar{x}) \geq 0$ for all $x \in C$, then $\bar{x}$ is a global minimizer of f in C.
- If in addition f is differentiable, then the sufficient condition becomes $-\nabla f(\bar{x}) \in N_C(\bar{x})$
- (in general, this statement does not hold for non-convex functions!)
- Proof:

Assume $-\nabla f(\bar{x}) \in N_c(\bar{x}) \Rightarrow -\nabla f(\bar{x})^T(x-\bar{x}) \leq 0 \quad \forall x \in C \Rightarrow f(\bar{x}) + \nabla f(\bar{x})^T(x-\bar{x}) \geq f(\bar{x}) \quad \forall x \in C$

$\Rightarrow$ since f is convex, $\underbrace{f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T(x-\bar{x})}_{\text{first-order condition}} \geq f(\bar{x}) \Rightarrow \bar{x}$ global min

More in general (if non differentiable), assume $f'(\bar{x}; x-\bar{x}) \geq 0 \quad \forall x \in C$

$\Rightarrow f(\theta x + \bar{\theta}\bar{x}) \leq \theta f(x) + \bar{\theta} f(\bar{x}) \Rightarrow \dfrac{f(\theta x + \bar{\theta}\bar{x}) - f(\bar{x})}{\theta} \leq f(x) - f(\bar{x})$

$\Rightarrow$ in the $\lim_{\theta \to 0}$: $\underbrace{f'(\bar{x}; x-\bar{x})}_{\geq 0} + f(\bar{x}) \leq f(x) \Rightarrow f(\bar{x}) \leq f(x) \quad \forall x \in C \Rightarrow \bar{x}$ global min

$\Rightarrow$ for f convex and differentiable, $\bar{x}$ global min $\Leftrightarrow -\nabla f(\bar{x}) \in N_c(\bar{x})$