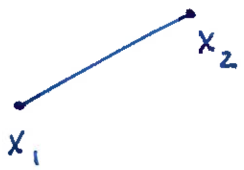
Elements of convex optimization

* Convex optimization problems (including linear programming problems) are rare exceptions of optimization problems that can be solved explicitly / efficiently.

* Problems of general form *might* be transformed into ~~optin~~ convex optimization problems.

$x_1, x_2 \in \mathbb{R}^n$



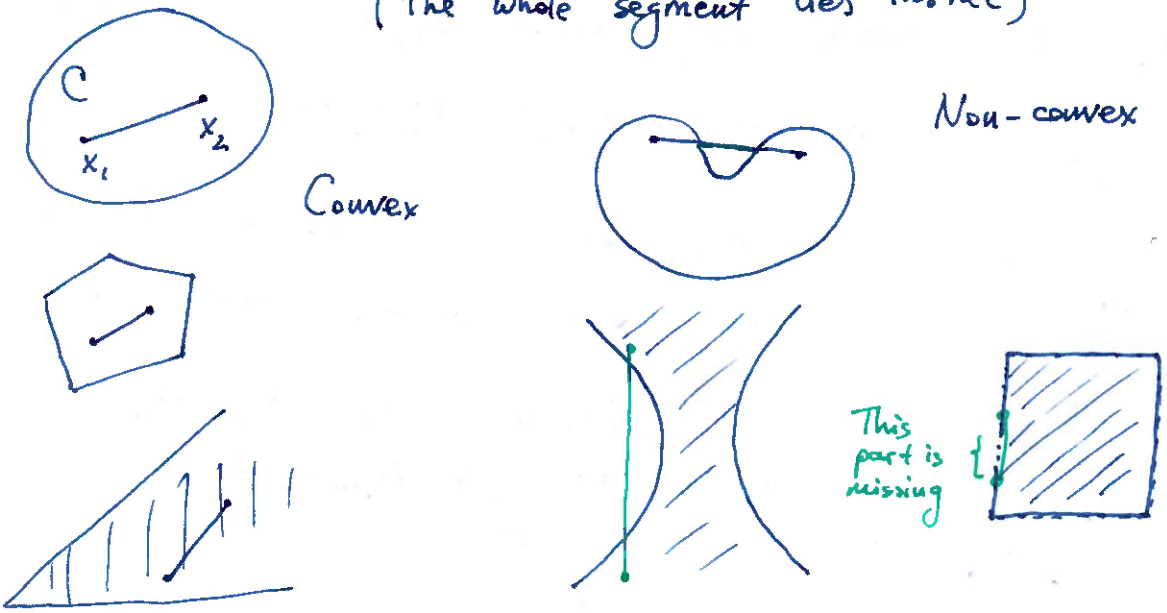
the line segment between x_1 and x_2 is the set

$$\{ x = \theta x_1 + (1-\theta) x_2 \mid \theta \in [0,1] \}$$

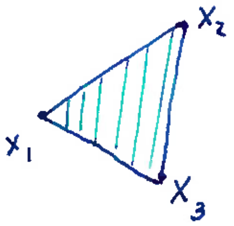
$C \subset \mathbb{R}^n$ is convex if

$$\forall x_1, x_2 \in C \quad \forall \theta \in [0,1], \quad \theta x_1 + (1-\theta) x_2 \in C$$

(the whole segment lies inside)



$x_1, x_2, \ldots, x_m$: the convex combination of these points is the set of all points of the form

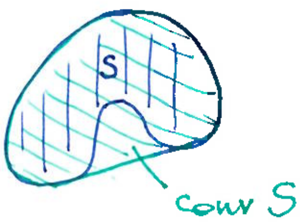


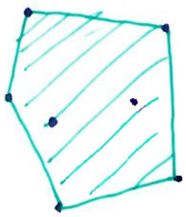
$$x = \theta_1 x_1 + \ldots + \theta_m x_k$$
$$\theta_1 + \ldots + \theta_k = 1, \quad \theta_i \geq 0$$

$S \subset \mathbb{R}^n$ set. The convex hull of S

$$\operatorname{conv} S = \{\text{all convex combinations of all possible points in } S\}$$

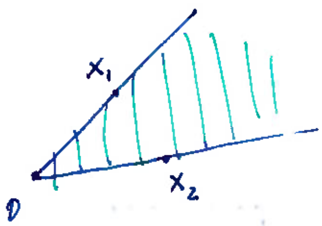
= The smallest (by inclusion) convex set that contains S





x_1, x_2 : convex cone over x_1, x_2 is the set of pt's

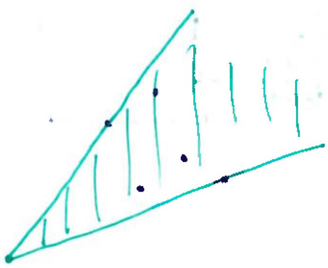
$$x = \theta_1 x_1 + \theta_2 x_2, \quad \theta_1, \theta_2 \geq 0$$

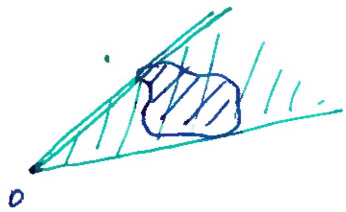


convex cone over set S is the set of all possible convex cone combinations

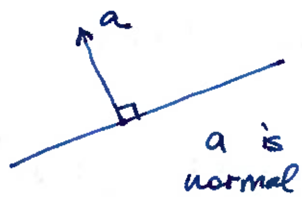
$$\theta_1 x_1 + \ldots + \theta_m x_m \quad \theta_1, \ldots \theta_m \geq 0$$

for possible pt's $x_1, \ldots x_m$ in S





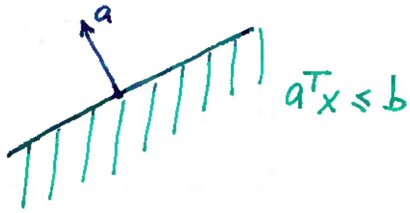
* $\{ x \in \mathbb{R}^n \mid a^T x = b \}$ <u>hyperplane</u>



a is the normal vector

$\{ x \in \mathbb{R}^n \mid a^T x \leq b \}$ <u>half space</u>

can be < 0, > 0, ≥ 0



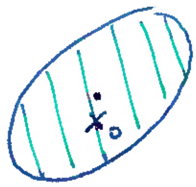
<u>Note</u>: hyperplanes / halfspaces are convex (Why?)

<u>Ex</u>: Check ~~by~~ using definition.

* Ellipsoids $\mathcal{E} = \{ x \in \mathbb{R}^n \mid (x - x_0)^T A (x - x_0) \leq 1 \}$ are convex

(here, A is symmetric and positive-definite.)

<u>Ex</u>: Show ~~by~~ using definition.



In particular, if $A = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$, then $\mathcal{E}$ is the <u>unit</u> ball centered at x_0.

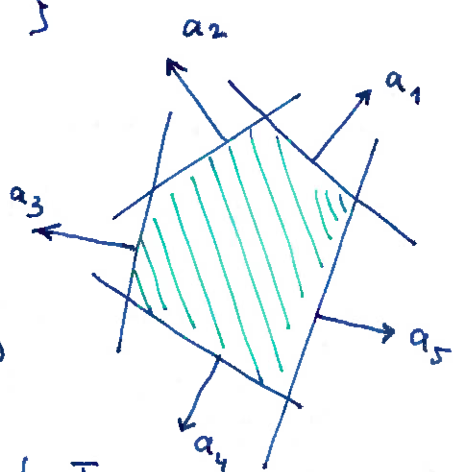
* Polyhedron is the solution of finitely many equalities and inequalities

$$\{ x \in \mathbb{R}^n : Ax \leq b , \quad Cx = d \}$$

$$m \Big\{ \underbrace{\begin{pmatrix} A \end{pmatrix}}_{n} \begin{pmatrix} \end{pmatrix} \Big\}^{n} \qquad p \Big\{ \underbrace{\begin{pmatrix} C \end{pmatrix}}_{n}$$

m inequalities (understood component-wise)

p equalities



$\{ a_i^T x \leq b_i, \ i = 1 \ldots 5$

<u>Lemma</u>: If C_1, C_2 convex, then $C_1 \cap C_2$ is convex $\Rightarrow$ <u>Cor.</u>: Polyhedra are convex

<u>Lemma</u>: <u>Any</u> intersection is convex (finite, countable, uncountable)

The last lemma allows to detect convex sets:

<u>Example:</u> $f: \mathbb{R}^n \to \mathbb{R}^m$ affine

$$f(x) = Ax + b .$$

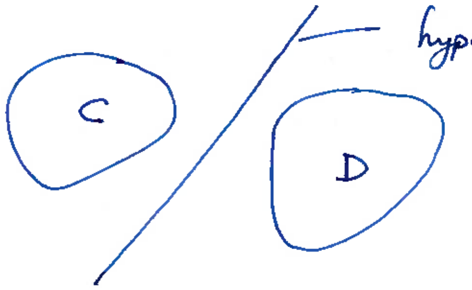
scalings, rotations, translations, flips, etc...

Then if $S \subset \mathbb{R}^n$ convex, then $f(S) = \{ f(x) \mid x \in S \}$ is also convex. Same for the inverse map f^{-1}

<u>Thm:</u> (<u>Separating hyperplane thm</u>): If $C, D \subset \mathbb{R}^n$ are convex and disjoint, then $\exists\ a \neq 0, b$ s.t.

$$a^T x \le b \quad \forall x \in C \quad , \quad a^T x \ge b \quad \forall x \in D.$$

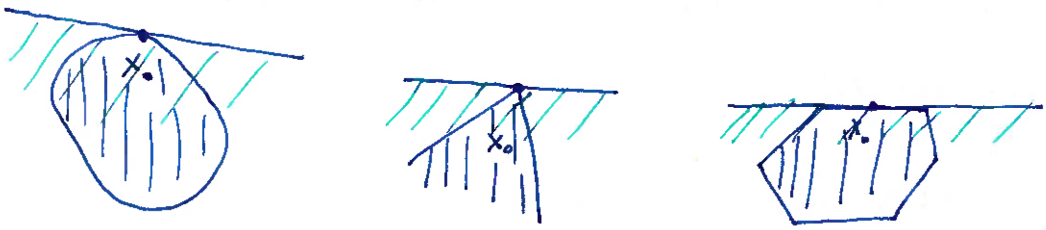
hyperplane $a^T x = b$ separates C and D.



? Strict separation: if C is ~~op~~ compact.

Supporting hyperplane: $x_0 \in \partial C$:

$\{ x : a^T x = a^T x_0 \}$ and $a^T x \le a^T x_0 \quad \forall x \in C$



Using Separating hyperplane thm: C convex, $\{x_0\}$ convex, $x_0 \in \partial C$

$\to\ \exists$ separating hyperplane

<u>Cor.:</u> Through every pt. on the boundary of a convex set passes (at least one) ~~sea~~ supporting hyperplane.

also convex

◀ $\operatorname{int} C \cap \{x_0\} = \emptyset$ ▶

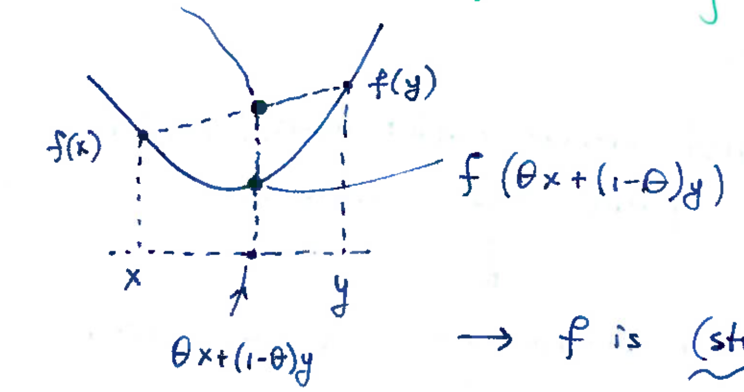
? What if $\operatorname{int} C = \emptyset$?

* Convex functions

$f : \mathbb{R}^n \to \mathbb{R}$ is convex if

(1) • domain of definition $\operatorname{dom} f$ is convex

(2) • $$f(\underbrace{\theta x + (1-\theta) y}_{\in \operatorname{dom} f \text{ since } \operatorname{dom} f \text{ is convex}}) \overset{*}{\leq} \theta f(x) + (1-\theta) f(y) \quad \forall x, y \in \operatorname{dom} f \quad \forall \theta \in [0,1]$$

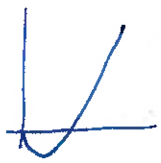


→ f is strictly convex if in $*$ we have $<$ for $\theta \in (0,1)$

→ f is (strictly) concave if $-f$ is (strictly) convex

Examples: of convex functions

$ax + b$, e^x, $|x|^p$ ($p \geq 1$), $\underbrace{x \log x}_{\text{negative entropy}}$ $\quad \mathbb{R}$



$a^T x + b$, $\|x\|_p = \left(\sum (x_i)^p\right)^{1/p}$, $\|x\|_\infty = \max_k x_k$ $\quad \mathbb{R}^n$

Lemma: $f : \mathbb{R}^n \to \mathbb{R}$ convex $\iff$ $g : \mathbb{R} \to \mathbb{R}$ convex

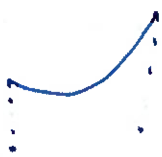
$g(t) = f(x + tv)$

$\operatorname{dom} g = \{ t \mid x + tv \in \operatorname{dom} f \}$

one variable! ⇓ easier

Note: Extended-value $\tilde{f}$ of f is .... $+\infty$

$$\tilde{f}(x) = \begin{cases} f(x), & x \in \operatorname{dom} f \\ +\infty, & x \notin \operatorname{dom} f \end{cases}$$



$\Rightarrow$ (1) & (2) $\iff$ $\tilde{f}(\theta x + (1-\theta) y) \leq \theta \tilde{f}(x) + (1-\theta) \tilde{f}(y)$

Just one condition!

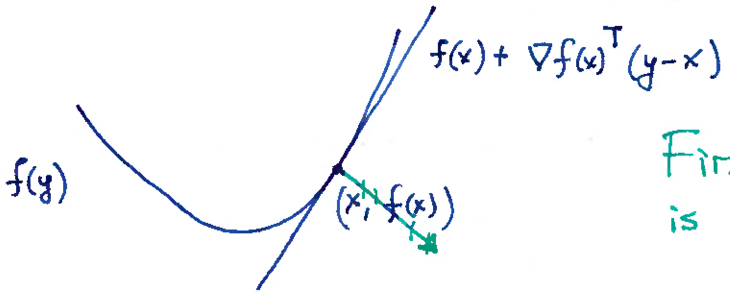
* Convex functions and ~~opt~~ differentiation

Assume f differentiable (smooth) on $\mathrm{dom}\, f$

<u>Thm</u>: (1^{st} order condition) f ~~differentiable~~ smooth on $\mathrm{dom}\, f$.

$\mathrm{dom}\, f$ is convex.

$$f \text{ convex} \iff f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall x, y \in \mathrm{dom}\, f$$



First order ~~estim~~ approximation is global underestimator
↓
local information ($\nabla f(x)$) gives global estimate

<u>Thm</u>: (2^{nd} order condition) f smooth on $\mathrm{dom}\, f$,

$\mathrm{dom}\, f$ is convex

$$f \text{ convex} \iff \nabla^2 f(x) \succeq 0 \quad \forall x \in \mathrm{dom}\, f$$

↑ positive semi-definite

If $\nabla^2 f(x) \succ 0$ $\forall x \in \mathrm{dom}\, f$, then f is strictly convex $\qquad f(x) = x^4$

<u>Examples</u>: ⊙ $f(x) = \frac{1}{2} x^T A x + b^T x + c$ quadratic function

$\nabla f(x) = Ax + b \qquad \nabla^2 f = A$

f is convex iff $A \succeq 0$.

⊙ least squares objective: $f(x) = \| Ax - b \|_2^2$

$\nabla f = 2A^T(Ax - b) \qquad \nabla^2 f = 2A^T A \succeq 0$ always $\Rightarrow$

$\Rightarrow$ f is convex

⊙ $f(x,y) = \frac{x^2}{y} \longrightarrow$ <u>Ex</u>: Convex

⊙ log - sum - exp: $f(x) = \log \sum_{k=1}^{n} e^{x_k}$ "soft max"

$$\nabla^2 f(x) = \frac{1}{\mathbb{1}^T z} \operatorname{diag}(z) - \frac{1}{(\mathbb{1}^T z)^2} z z^T \qquad z_k = e^{x_k} \qquad z = \begin{pmatrix} e^{x_1} \\ e^{x_2} \\ \vdots \\ e^{x_n} \end{pmatrix}$$

positive semidef

$$\rightarrow \quad v^T \nabla^2 f(x) v = \frac{\left(\sum z_k v_k^2\right)\left(\sum z_k\right) - \left(\sum z_k v_k\right)^2}{\left(\sum z_k\right)^2} \geq 0$$

Cauchy-Schwarz

⊙ geometric mean: $f(x) = -\sqrt[n]{x_1 x_2 \cdot \ldots \cdot x_n}$

Thm: f convex function, $\operatorname{dom} f$ conv. Then every loc. minimum x^* is a global minimum

◀ $f(x) \geq f(x^*)$ (#) $\forall x \in B(x^*, \varepsilon) \cap \operatorname{dom} f$.

Take $z \in \operatorname{dom} f$. Then $(1-\theta) x^* + \theta z \in B(x^*, \varepsilon) \cap \operatorname{dom} f$ for small θ $\Rightarrow$

$\overset{(\#)}{\Rightarrow} f((1-\theta)x^* + \theta z) \geq f(x^*)$

$(1-\theta) f(x^*) + \theta f(z) \geq f((1-\theta)x^* + \theta z)$ (convexity)

$\Rightarrow (1-\theta) f(x^*) + \theta f(z) \geq f(x^*) \Rightarrow \theta f(z) \geq \theta f(x^*)$

$\Rightarrow f(z) \geq f(x^*)$ ▶

Thm: f convex, $\operatorname{dom} f$ convex.

$C_\alpha = \{ x \in \operatorname{dom} f \mid f(x) \leq \alpha \}$ is convex

In particular C_0 is convex $\Rightarrow$ 1^{st} order condition

using supporting hyperplane theorem

Rem: ∇f is a normal to the level lines.

Cor.: x^* global min. $\Leftrightarrow \nabla f(x^*) = 0$

Epigraph $\operatorname{epi}(f) = \{ (x,t) \in \mathbb{R}^{n+1} \mid x \in \operatorname{dom} f,\ f(x) \leq t \}$

__Prop.:__ $\operatorname{epi}(f)$ convex $\Leftrightarrow$ f is convex.

* What preserve convexity:

__Prop:__ f convex $\quad f_1, f_2$ convex

1) αf convex $\forall \alpha$

2) $f_1 + f_2$ convex (infinite sums)

3) $f(Ax+b)$ $\longrightarrow$ __Ex:__ $f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x)$

barrier function

$a_i^T x \leq b_i$

4) $f(x) = \max\{f_1(x) \dots f_m(x)\}$ convex

$\operatorname{epi}(\max f_i) = \bigcap_i \operatorname{epi}(f_i)$

5) $f(x,y)$ convex in x $\forall y \in \mathcal{A}$

$g(x) = \sup_{y \in \mathcal{A}} f(x,y)$ $\longrightarrow$ __Ex.:__

is convex

$f(x) = \sup_{y \in C} \|x-y\|$

__Ex:?__ $g: \mathbb{R}^n \to \mathbb{R}$, $h: \mathbb{R} \to \mathbb{R}$

When ~~does~~ is $f(x) = h(g(x))$ convex?

(assuming g convex)