L9 - 15/12/23 - Optimization.

Elements of convex optimization

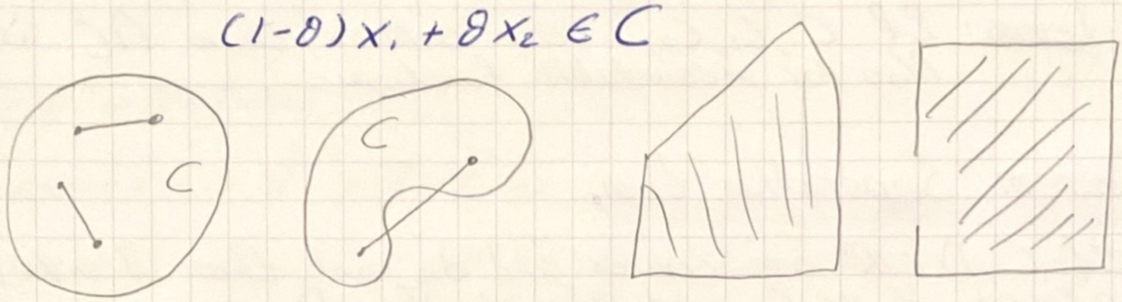
(1) Convex sets / convex functions

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

$\{x \in \mathbb{R}^n : x = (1-\theta)x_1 + \theta x_2 \text{ for } \theta \in [0,1]\}$

$C \subset \mathbb{R}^n$ is a convex set if $\forall\, x_1, x_2 \in C$, $\forall\, \theta \in [0,1]$

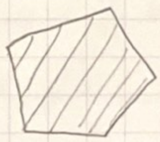
$$(1-\theta)x_1 + \theta x_2 \in C$$



convex — not convex — convex — not convex

$x_1, \dots, x_m \in \mathbb{R}^n$, the convex combination of $x_1, \dots, x_m$ is the set

$$\theta_1 x_1 + \theta_2 x_2 + \dots + \theta_m x_m, \quad \theta_i \geq 0, \quad \theta_1 + \theta_2 + \dots + \theta_m = 1$$



$x_1, x_2, x_3 \longrightarrow S \subset \mathbb{R}^n$

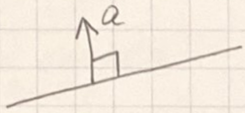
$\theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$

conv S = smallest convex set that contains S
= all possible convex combinations of points in S

↑ convex hull

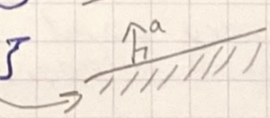
Key examples:

- hyperplane $\{x : a^T x = b\}$



- half space $\{x, a^T x \leq b\}$

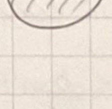
convex sets



- ellipsoids $\mathcal{E} = \{x \in \mathbb{R}^n : (x - x_0)^T A (x - x_0) \leq 1\}$

A symmetric positive-definite matrix

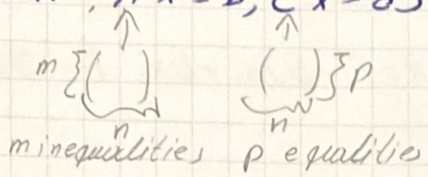
In particular, if $A = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ ball $\|x - x_0\|_2^2 \leq 1$ of radius 1 is also convex



→

- Polyhedion is the solution of finitely many linear equalities/ inequalities

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



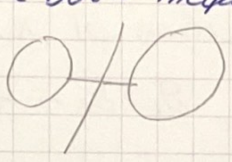
- Polyhedra are convex sets. This follows from

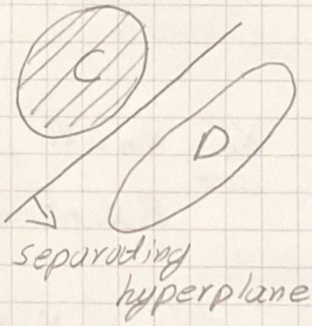
  Lemma: If $C_1, C_2, C_3, \dots$ are convex, then $\bigcap C_i$ is convex. Also for uncountable functions

Theorem: Seperation thm

If $C, D \subset \mathbb{R}^n$ are convex and disjoint, then $\exists\, a \neq 0$, b s.t. $a^T x \leq b \;\forall\, x \in C$ and $a^T x \geq b \;\forall\, x \in D$

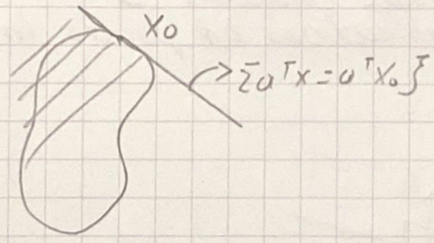
If C is compact then there is a strict separation (both inequalities are strict)





Supporting hyperplane of a set C

$x_0 \in \partial C$ if $\exists\, a$ s.t. $a^T x \leq a^T x_0 \;\forall\, x \in C$

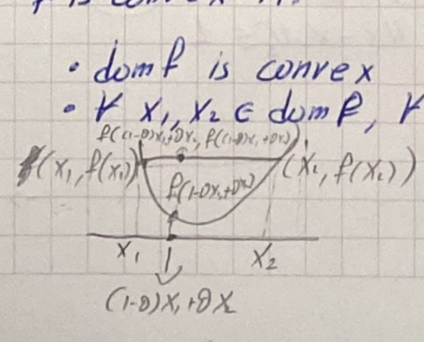


Cor: Through every point at the boundary of a convex set passes a supporting hyperplane

Convex functions: $f: \mathbb{R}^n \to \mathbb{R}$ (dom f domain of f)

f is convex if:

- dom f is convex
- $\forall\, x_1, x_2 \in \operatorname{dom} f$, $\forall\, \theta \in [0,1]$, $f((1-\theta)x_1 + \theta x_2) \overset{(*)}{\leq} (1-\theta) f(x_1) + \theta f(x_2)$



→ f is strictly convex if $(*)$ is $<$

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

Examples: ($\mathbb{R}$) $ax+b$, e^x, $|x|^p$ $p>1$, $x\log x$

($\mathbb{R}^n$) a^Tx+b, $\|x\|_p = \sqrt[p]{\sum |x_i|^p}$, $p\geq 1$

$\|x\|_\infty = \max_k |x_k|$

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

$g(t)=f(x+tv)$
$\operatorname{dom} g = \{t : x+tv \in \operatorname{dom} f\}$

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

$\tilde{f}:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$

f is convex $\iff$ $\tilde{f}$ satisfies

$$\tilde{f}((1-\theta)x_1+\theta x_2) \leq (1-\theta)\tilde{f}(x_1)+\theta\tilde{f}(x_2) \quad \forall x_1, x_2 \in \mathbb{R}^n, \ \forall \theta\in[0,1]$$

(2) Convexity and optimality

Theorem: f convex

$C_a = \{x\in \operatorname{dom} f : f(x)\leq a\}$ is convex

▶ $x, y \in C_a$, $f((1-\theta)x+\theta y) \leq (1-\theta)f(x)+\theta f(y) \leq (1-\theta)a+\theta a = a$

$\Rightarrow (1-\theta)x+\theta y \in C_a$ ◼

In particular C_0 is convex

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

Theorem: f is convex $\iff$ $\operatorname{epi}(f)$ convex
(proof using the idea above)

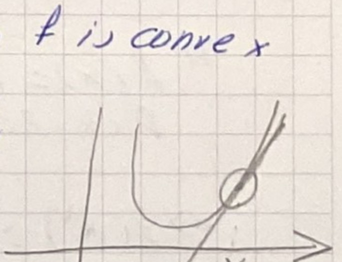
$\Rightarrow \forall x \in \partial\operatorname{epi}(f)$ there exists a supporting hyperplane

Assume f is smooth $\rightarrow$ supporting hyperplane is tangent

Theorem: (1st order condition) f smooth, $\operatorname{dom} f$ is convex

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

$\uparrow$
from local data ($\nabla f(x)$)
$\rightsquigarrow$
global condition



Theorem (2nd order condition): f smooth, dom f convex
(positive semi-definete) $\forall x \in \text{dom} f$

$$f \text{ convex} \iff \nabla^2 f(x) \geq 0$$

If $\nabla f(x) > 0$, then f is strictly convex

Examples:

(1) $f(x) = \frac{1}{2} x^T A x + b^T x + c$ quadratic function

$\nabla f(x) = Ax + b$, $\nabla^2 f(x) = A$ convex $\iff A \geq 0$

(2) least squares objective:

$f(x) = \|Ax - b\|_2^2$ → squared, ↳ euclidean

$\nabla f(x) = 2A^T(Ax - b)$

$\nabla^2 f(x) = 2A^T A$ ← positive semi-definite

(3) Soft maximum

$f(x) = \log\left(\sum_{i=1}^{n} e^{x_i}\right)$ $x = (x_1, \dots, x_n)$

This is a convex function; $z = (e^{x_1}, e^{x_2}, \dots, e^{x_n})$

$$\text{diag}(z) = \begin{pmatrix} z_1 & & 0 \\ & z_2 \dots & \\ 0 & & z_n \end{pmatrix}$$

claim: $\nabla^2 f(x) = \frac{1}{z_1 + z_2 + \dots + z_n} \text{diag}(z) = -\frac{1}{(z_1 + \dots + z_n)^2} z z^T$

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

by Cauch-Schwaz Ineq.

(4) geometric mean $x_i \geq 0$ $f(x) = \sqrt[n]{x_1 \dots x_n}$ convex function

Theorem: f convex function. Then every local minimum x^* is to a local minimum

□ $f(x) \geq f(x^*)$ $\forall x \in B(x; \varepsilon) \cap \text{dom} f$

Pick $z \in \text{dom} f$, $(1-\theta)x^* + \theta z \in \text{dom} f \cap B(x^*, \varepsilon)$

$\theta \in [0, 1]$ ↳ choose θ sufficiently small

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

$\Rightarrow \cancel{\theta} f(x^*) \leq \cancel{\theta} f(z)$ ■

Cor: x^* global minimum $\Leftrightarrow \nabla f(x^*) = 0$
(if f is smooth)

(3) Operations w/ convex functions

Prop: (1) f convex, $a \geq 0$, then f is convex

(2) If f_1, f_2 convex, then $f_1 + f_2$ is convex

(3) If $f(x)$ convex, then $f(Ax+b)$ convex
(Example: barrier functions, $f(x) = -\sum_{i=1}^{n} \log(b_i - a_i^T x)$)

(4) $f(x) = \max_k \{f_1(x) \ldots f_k(x)\}$, f_i convex.

(5) $f(x,y)$ is convex in x for any $y \in A$, then

$g(x) = \sup_{y \in A} f(x,y)$ is convex