Convex Optimisation

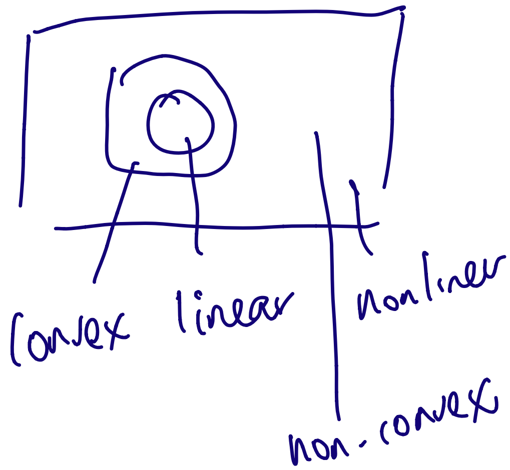
Matthew Tam
(University of Melbourne)
matthew.tam@unimelb.edu.au

ACE Network
S1 2021

Course webpage:

matthewktam.github.io/cvxopt

2.1 Convexity

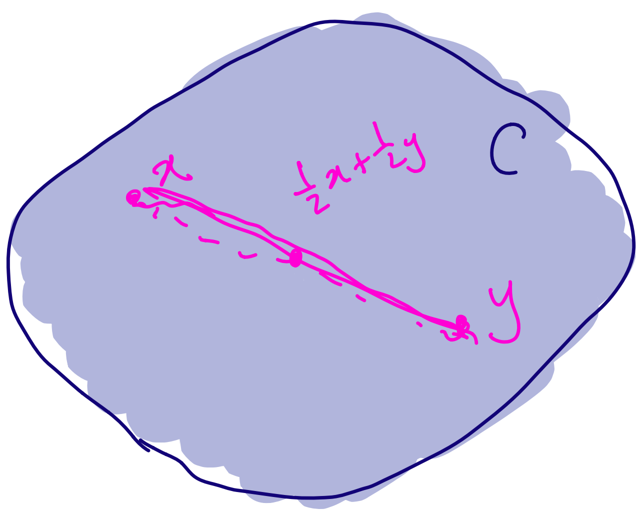


Definition 2.1.1

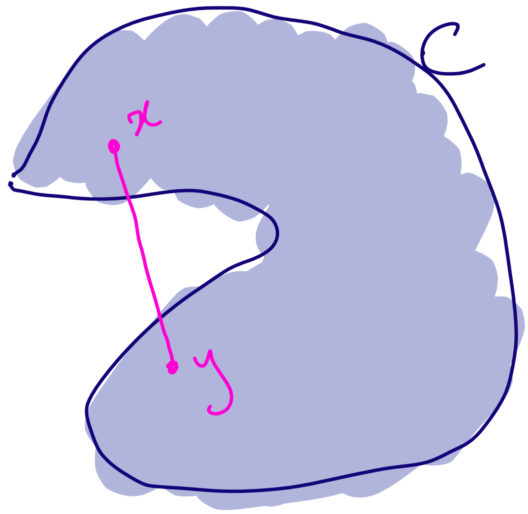
A set $C \subseteq \mathbb{R}^n$ is convex if

$$x, y \in C \Rightarrow \lambda x + (1-\lambda) y \in C \quad \forall \lambda \in [0,1]$$

Eg.



C is a convex set



C is not a convex set

Examples of convex sets

(a) The empty set $\emptyset$, and entire space $\mathbb{R}^n$

(b) Let $a \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}$.

(i) The <u>hyperplane</u> given by

$$\{x \in \mathbb{R}^n : \langle a, x \rangle = b\}$$

(ii) The <u>half-space</u> given by

$$H = \{x \in \mathbb{R}^n : \langle a, x \rangle \leq b\}$$

Pf Let $x, y \in H$ and $\lambda \in [0,1]$. Then:

$$\langle a, \lambda x + (1-\lambda) y \rangle$$

$$= \lambda \langle a, x \rangle + (1-\lambda) \langle a, y \rangle$$

$$\leq \lambda b + (1-\lambda) b \qquad [\text{Note } \lambda, 1-\lambda \geq 0].$$

$$= b$$

i.e. $\lambda x + (1-\lambda) y \in C$.

(c) Balls: $B_\delta(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\| \leq \delta\}$

where $x_0 \in \mathbb{R}^n$ and $\delta > 0$.

Pf Let $x, y \in B_\delta(x_0)$ and $\lambda \in [0,1]$. Then:

$$\| \lambda x + (1-\lambda) y - x_0 \|$$

$$= \| \lambda (x - x_0) + (1-\lambda)(y - x_0) \|$$

$$\leq \| \lambda (x - x_0) \| + \| (1-\lambda)(y - x_0) \| \qquad (\Delta\text{-ineq})$$

$$= \lambda \| x - x_0 \| + (1-\lambda) \| y - x_0 \|$$

$\leq \lambda\delta + (1-\lambda)\delta = \delta.$

i.e. $\lambda x + (1-\lambda)y \in B_\delta(x_0)$.

(d) Intersection of convex sets.

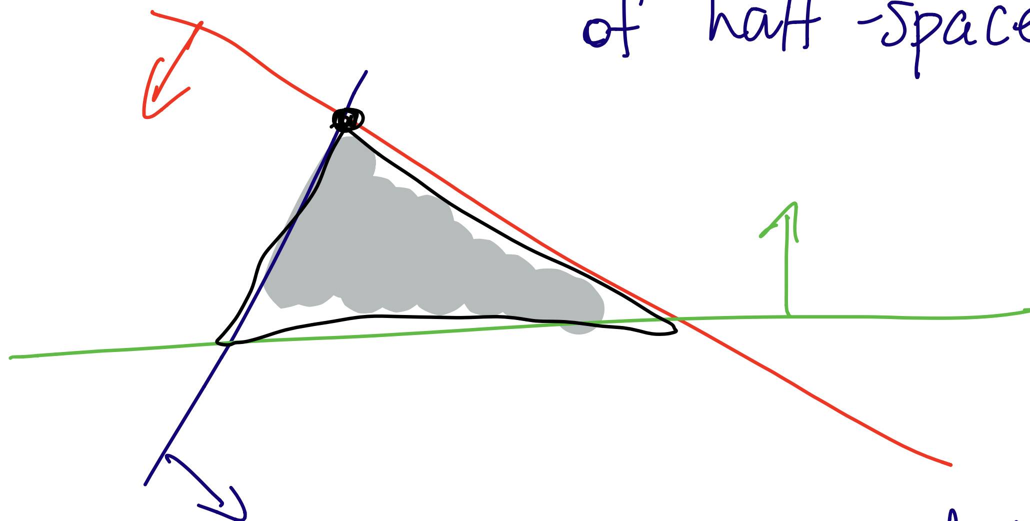
Let $C_i \subseteq \mathbb{R}^n$ be convex for all $i \in I$, and denote $C := \bigcap_{i\in I} C_i$.
Then C is convex.

Pf. Let $x, y \in C$ and $\lambda \in [0,1]$.
Then $x, y \in C_i$ for all $i \in I$.
Since C_i is convex, it follows that

$$\lambda x + (1-\lambda)y \in C_i \quad \forall i \in I.$$

$$\Rightarrow \lambda x + (1-\lambda)y \in C.$$

(e) Polyhedra (= finite intersection of half-spaces)



((Convex by combining (b) and (d))).

In convex analysis, we allow <u>extended real-valued</u> functions. That is, they can take the values $\pm\infty$. When working with infinities, we adopt the following conventions:

$$+\infty + \infty = +\infty.$$
$$-\infty - \infty = -\infty$$
$$\pm\infty \mp \infty = \text{undefined.}$$

$$+\infty + x = +\infty \quad \forall x \in \mathbb{R}.$$
$$-\infty + x = -\infty \quad \forall x \in \mathbb{R}.$$

(all operations commute)

<u>From lecture 1</u>:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad x \in C. \qquad (1.1)$$

where $f : \mathbb{R}^n \to \mathbb{R}$.

This can be formulated as

$$\min_{x \in \mathbb{R}^n} \tilde{f}(x) \quad \text{where } \hat{f}(x) = \begin{cases} f(x), & x \in C \\ +\infty, & x \notin C. \end{cases}$$

Defintion 2.1.2 (convex function)

A function $f: \mathbb{R}^n \to (-\infty, +\infty]$ is <u>convex</u> if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

for all $x, y \in \operatorname{dom} f := \{x \in \mathbb{R}^n : f(x) < +\infty\}$ and $\lambda \in [0,1]$. If the inequality is strict whenever $\lambda \in (0,1)$ and $x \neq y$, then we say f is <u>strictly convex</u>.

<u>Remark</u>: We will always deal with convexity, but one can also work with <u>concave functions</u>. A f is concave if $-f$ is convex.

Example 2.1.3

(a) Every linear (or affine) is convex.
(but not strictly convex).

Pf Let $f:\mathbb{R}^n \to \mathbb{R}$ be linear, $x,y\in\mathbb{R}^n$ and $\lambda\in[0,1]$. Then

$$f(\lambda x+(1-\lambda)y) = \lambda f(x)+(1-\lambda)f(y)$$
$$\le \lambda f(x)+(1-\lambda)f(y)$$

(b) Any norm is a convex function.

Pf. Let $p:\mathbb{R}^n \to \mathbb{R}_+$ be a norm.
Take $x,y\in\mathbb{R}^n$ and $\lambda\in[0,1]$. Then:

$$\begin{aligned} & p(\lambda x+(1-\lambda)y) \\ & \le p(\lambda x)+p((1-\lambda)y) && (\Delta\text{-ineq.}) \\ & = \lambda p(x)+(1-\lambda)p(y) && (\text{pos. homo.}) \end{aligned}$$

ie. p is convex

We say $p : \mathbb{R}^n \to \mathbb{R}_+$ is a <u>norm</u> if.

(a) $p(x+y) \leq p(x) + p(y) \quad \forall x, y \in \mathbb{R}^n$

(b) $p(\lambda x) = |\lambda| p(x) \quad \forall x \in \mathbb{R}^n, \lambda \in \mathbb{R}^n$

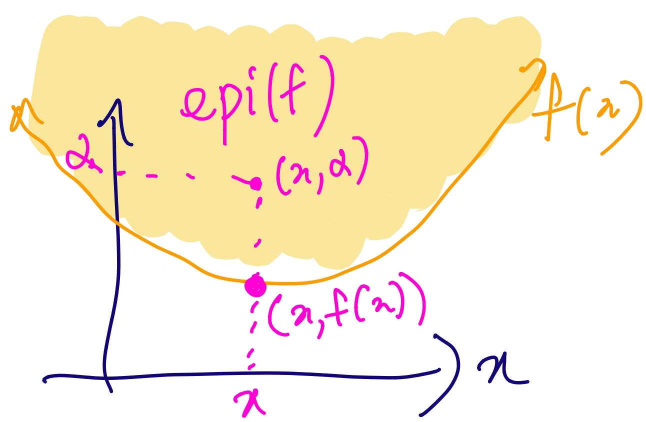
(c) $p(x) = 0 \iff x = 0$.

Exercise 2.1.4

Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a function, and let $C \subseteq \mathbb{R}^n$ be a set.

(a) f is convex if and only if its epigraph is a convex set.

$$\text{epi}(f) := \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq \alpha\}.$$



(b) If f is convex, then dom f is a convex set.

(c) C is convex if and only if its indicator function $i_C : \mathbb{R}^n \to (-\infty, +\infty]$ is a convex function, where

$$i_C(x) := \begin{cases} 0 & , \quad x \in C \\ +\infty & , \quad x \notin C \end{cases}$$

Proprosition 2.1.5

(a) Let $f, g : \mathbb{R}^n \to (-\infty, +\infty]$ and $\alpha \geq 0$. Then $f + \alpha g$ is convex.

(b) Let $f_i : \mathbb{R}^n \to (-\infty, +\infty]$ for all $i \in I$. Then $f(x) := \max_{i \in I} f_i(x)$ is convex.

(c) Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be linear/affine and $f : \mathbb{R}^m \to (-\infty, +\infty]$ be convex. Then $f \circ A$ is convex.

<u>Pf</u> (a) Let $x, y \in \operatorname{dom}(f+\alpha g) = \operatorname{dom} f \cap \operatorname{dom} g$, and $\lambda \in [0,1]$. Using convexity of f and g, we deduce

$$
\begin{aligned}
(f+\alpha g)(\lambda x + (1-\lambda)y) &= f(\lambda x + (1-\lambda)y) + \alpha g(\lambda x + (1-\lambda)y) \\
&\leq \lambda f(x) + (1-\lambda) f(y) \\
&\qquad + \alpha\left[\lambda g(x) + (1-\lambda) g(y)\right] \\
&= \lambda (f+\alpha g)(x) + (1-\lambda)(f+\alpha g)(y)
\end{aligned}
$$

This completes the proof. $\blacksquare$

$$
\begin{aligned}
\operatorname{dom}(f+\alpha g) &= \{x \in \mathbb{R}^n : f(x) + \alpha g(x) < +\infty\} \\
&= \{x : f(x) < +\infty\} \\
&\qquad \cap \{x : g(x) < +\infty\} \\
&= \operatorname{dom} f \cap \operatorname{dom} g.
\end{aligned}
$$

Proposition 2.1.6

Let $f: \mathbb{R}^n \to (-\infty, +\infty]$ be convex.

Then

$$\underset{x \in \mathbb{R}^n}{\operatorname{argmin}} f(x) := \left\{ x \in \mathbb{R}^n : f(x) = \inf_{y \in \mathbb{R}^n} f(y) \right\}$$

is a convex set. Furthermore, if f is strictly convex, $\underset{x \in \mathbb{R}^n}{\operatorname{argmin}} f(x)$ is either empty or singleton.

Pf. Let $x, y \in \underset{z \in \mathbb{R}^n}{\operatorname{argmin}} f(z)$ and $\lambda \in [0,1]$.
Further, denote $f^* = \min_{z \in \mathbb{R}^n} f(z) = f(x) = f(y)$.
By convexity of f, we have

$$\begin{aligned} f^* &\le f(\lambda x + (1-\lambda) y) \\ &\le \lambda f(x) + (1-\lambda) f(y) \end{aligned}$$

$$= \lambda f^* + (1-\lambda) f^*$$
$$= f^*.$$

That is,

$$f^* \le f(\lambda x + (1-\lambda)y) \le f^*$$
$$\Rightarrow f(\lambda x + (1-\lambda)y) = f^*$$
$$\Rightarrow \lambda x + (1-\lambda)y \in \operatorname*{argmin}_{z \in \mathbb{R}^n} f(z).$$

This the first part of the proposition.

Now, assume f is strictly convex.
Take $x, y \in \operatorname*{argmin}_{z \in \mathbb{R}^n} f(z)$ with $x \neq y$.
Then repeat the above arguement with $\lambda \in (0,1)$ to obtain

$$f^* \le f(\lambda x + (1-\lambda)y)$$
$$< \lambda f(x) + (1-\lambda) f(y)$$

$$= \lambda f^* + (1-\lambda) f^*$$
$$= f^*.$$

That is,

$$f^* < f^*$$

which is a contradiction. Hence either $\operatorname{argmin}_{z \in \mathbb{R}^n} f(z) = \emptyset$ or $x = y$.

(i.e. $\operatorname{argmin}_{z \in \mathbb{R}^n} f(z)$ is a singleton).

$\blacksquare$