Convex Optimisation

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

ACE Network
S1 2021

Mid-semester break next week

* No lecture on April 8th
* Next lecture on April 15th.
  Assignment 2 due.

Last time: proximity operator

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

$$\operatorname{prox}_f(x) = \underset{y \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ f(y) + \frac{1}{2}\|x-y\|^2 \right\}$$

Th 2.3.4 Let f be proper, lsc and convex. Then prox_f is well-defined. Moreover, $p = \operatorname{prox}_f(x)$ iff

$$f(p) + \langle y-p, x-p \rangle \le f(y) \quad \forall y \in \mathbb{R}^n$$

Corollary 2.3.8

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

$$\left\langle \operatorname{prox}_f(x) - \operatorname{prox}_f(y), (I - \operatorname{prox}_f)(x) - (I - \operatorname{prox}_f)(y) \right\rangle \geq 0 \quad \forall x, y \in \mathbb{R}^n$$

or, equivalently,

$$\|\mathrm{prox}_f(x) - \mathrm{prox}_f(y)\|^2 + \|(I-\mathrm{prox}_f)(x) - (I-\mathrm{prox}_f)(y)\|^2 \le \|x-y\|^2 \quad \forall x,y \in \mathbb{R}^n.$$

In particular, prox_f is 1-Lipschitz continuous.

<u>Proof</u>: Exercise. (using Th 2.3.4). ⊠

C is nonempty, closed, convex subset of $\mathbb{R}^n$

$$\mathrm{prox}_{i_C}(x) = \underset{y\in\mathbb{R}^n}{\mathrm{argmin}}\left\{ i_C(y) + \tfrac{1}{2}\|x-y\|^2 \right\}$$

$$= \underset{y\in C}{\mathrm{argmin}}\left\{ \tfrac{1}{2}\|x-y\|^2 \right\}$$

$$= \underset{y\in C}{\mathrm{argmin}}\ \|x-y\| =: P_C(x)$$

Corollary 2.3.6

Let $C \subseteq \mathbb{R}^n$ be nonempty, closed and convex. Then P_C is well-defined. Moreover,

$p_k = P_C(x_k)$ iff $p \in C$ and

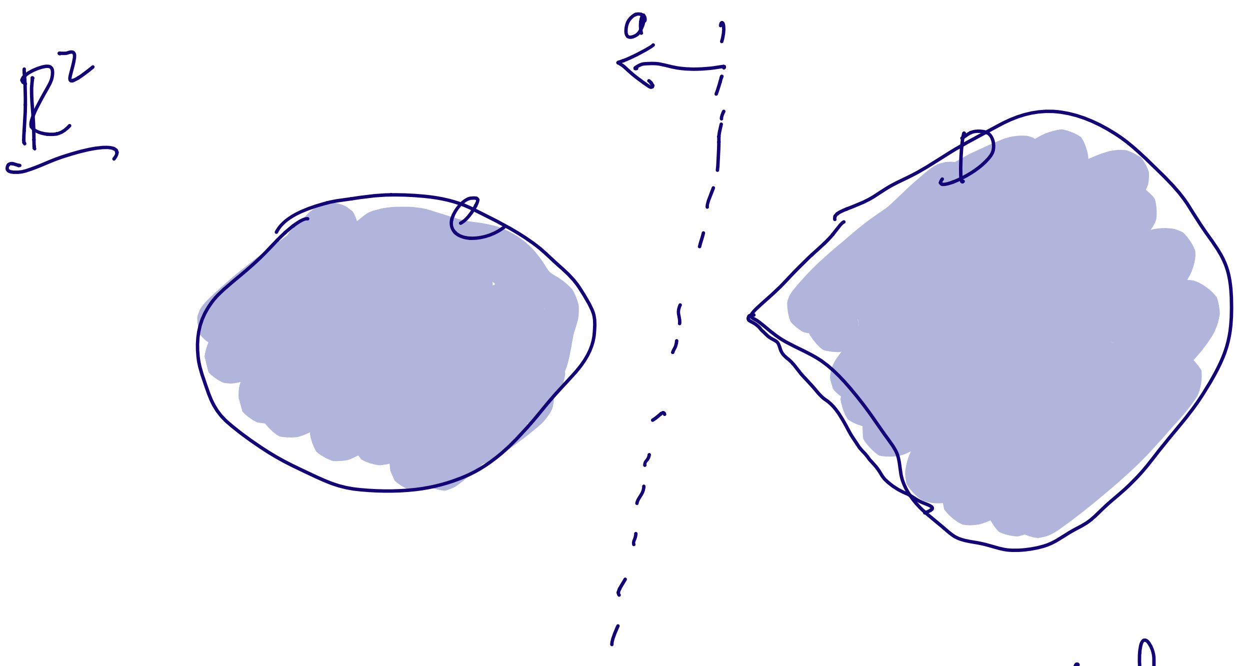
$$\langle x_k - p_k, z - p_k \rangle \leq 0 \quad \forall z \in C. \quad (*)$$

2.4 Separation theorems

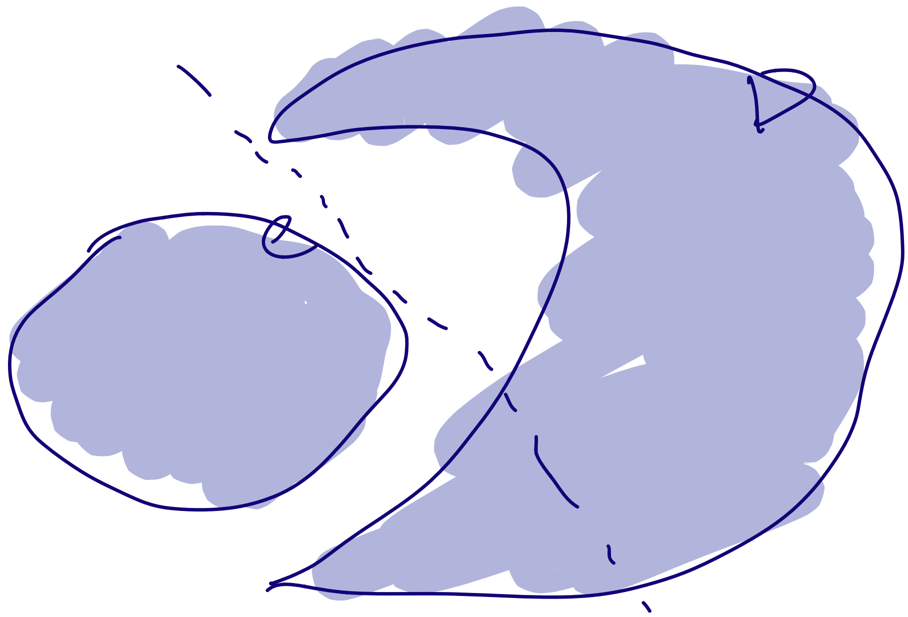
We consider conditions under which two set can be "separated" by a hyperplane. A hyperplane H is a set of the form

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

for some $a \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}$.



Two convex sets that can be separated.



Two sets that cannot be 'separated' by a hyperplane.

Theorem 2.4.1 (supporting hyperplane theorem)
Let $C \subseteq \mathbb{R}^n$ be a nonempty, convex set, and let $x_0 \in \mathbb{R}^n$ such that $x_0 \in \text{bdry}\, C$ or $x_0 \notin C$. Then there exists $a \in \mathbb{R}^n \setminus \{0\}$ such that

$$\langle a, z \rangle \leq \langle a, x_0 \rangle \quad \forall z \in C$$

Recall

$$cl(C) = \text{bdry}\, C \cup \text{int}\, C.$$

and

$$C \subseteq cl(C)$$

Proof. It suffices to establish the result for closed sets C. In this case, we consider $x_0 \notin C$. Let $(x_k) \notin C$ such that $x_k \to x_0$. Denote $p_k = P_C(x_k)$ and consider

$$a_k := \frac{x_k - p_k}{\|x_k - p_k\|} \qquad \forall k \geq 1.$$

Then (a_k) is bounded, hence it contain a subsequence which converges to a point $a \in \mathbb{R}^n$ (Th 1.1.6).

Since $p_k = P_C(x_k)$, Corollary 2.3.6 implies

$$\langle x_k - p_k, z - p_k \rangle \le 0 \quad \forall z \in C.$$

$$\Leftrightarrow \frac{1}{\|x_k - p_k\|} \langle x_k - p_k, z - p_k \rangle \le 0 \quad \forall z \in C.$$

$$\Leftrightarrow \quad \langle a_k, z - p_k \rangle \le 0 \quad \forall z \in C.$$

$$\Leftrightarrow \quad \langle a_k, z \rangle \le \langle a_k, p_k \rangle \quad \forall z \in C. \quad (*)$$

Next, note that

$$\langle a_k, p_k \rangle = \langle a_k, p_k - x_k \rangle + \langle a_k, x_k \rangle$$

$$= \left\langle \frac{x_k - p_k}{\|x_k - p_k\|}, p_k - x_k \right\rangle + \langle a_k, x_k \rangle$$

$$= -\frac{\|x_k - p_k\|^2}{\|x_k - p_k\|} + \langle a_k, x_k \rangle$$

$$\le \langle a_k, x_k \rangle \quad (**).$$

Now, combining (**) and (***) gives,

$$\langle \underset{\substack{\downarrow\\ a}}{a_n}, z\rangle \le \langle \underset{\substack{\downarrow\\ a}}{a_n}, \underset{\substack{\downarrow\\ x_0}}{x_n}\rangle \quad \forall z \in C.$$

Taking the limit along the sequence of (a_n) which converges to a gives

$$\langle a, z\rangle \le \langle a, x_0\rangle \quad \forall z \in C.$$

Which proves the result. $\blacksquare$

__Corollary 2.4.2__ (separating hyperplane th).

Let $C, D \subseteq \mathbb{R}^n$ be nonempty, convex sets with $C \cap D = \phi$. Then there exists $a \in \mathbb{R}^n \setminus \{0\}$ such

$$\langle a, x \rangle \leq \langle a, y \rangle \qquad \forall x \in C, y \in D.$$

__Proof__. Let $F \subseteq \mathbb{R}^n$ be defined as

$$F := C - D = \{x - y \in \mathbb{R}^n : x \in C, y \in D\}.$$

Then F is a nonempty convex set, and $0 \notin F$ as $C \cap D = \phi$. By Th 2.4.1,
x_0
there exists $a \in \mathbb{R}^n \setminus \{0\}$ such that

$$\langle a, z \rangle \leq \langle a, x_0 \rangle = 0 \qquad \forall z \in F$$

$$\iff \langle a, x-y \rangle \leq 0 \quad \forall x \in C, y \in D.$$

$$\iff \langle a, x \rangle - \langle a, y \rangle \leq 0 \quad \forall x \in C, y \in D$$

from which the result follows $\square$