2 Lecture 2

2.1 Topological Properties of Convexity

We are now going to move onto discuss the topological properties of convex sets. If you are awake, your immediate reaction should be 'with respect to what topology'? So far, we have only discussed convex sets in the context of a linear space, and we have (in general) defined a topology through the concept of a metric space. So what are we going to do?

In fact, there are two ways that we can go here. One is that we can restrict ourselves to normed linear spaces (or even $\mathbb{R}^n$) and use the topology inherent in such spaces (i.e. the topology generated by the metric associated with that norm). Another would be to use the algebraic structure of convex sets to define a new topology. Due to time constrains, we are going to do the former, though you should be aware of the latter possibility.

The properties that we are going to show may seem a little random, but they will turn out to be useful when we start separating things later.

Our first result is going to allow us to extend the convexity property to the closure of a set

Lemma 2 *Let C be a convex set in some metric space M. Let $x_1 \in int(C)$ and $x_2 \in cl(C)$. Then $[x_1, x_2) \subset int(c)$*

Proof. *First assume that $x_2 \in C$. The fact that $x_1 \in C^o \Rightarrow \exists\ r > 0$ such that $B(x_1, r) \subset C$. Let $z \in [x_1, x_2)$, then for some $\lambda \in (0, 1]$*

$$z = \lambda x_1 + (1 - \lambda)x_2$$

Now, say $y \in B(z, \lambda r)$. then

$$\begin{aligned} y &= z + (y - z) = z + v \\ &= \lambda x_1 + (1 - \lambda)x_2 + v \\ &= \lambda(x_1 + \frac{1}{\lambda}v) + (1 - \lambda)x_2 \end{aligned}$$

but

$$\begin{aligned} & d(x_1, x_1 + \frac{1}{\lambda}v) \\ = & \ ||x_1 - x_1 - \frac{1}{\lambda}v|| \\ = & \ ||\frac{1}{\lambda}v|| \\ = & \ \frac{1}{\lambda}||v|| \end{aligned}$$

and, as $v = (y - z)$*, then* $||v|| = d(y, z) \leq \lambda r$*, so*

$$\begin{aligned} x_1 + \frac{1}{\lambda}v & \in B(x_1, r) \subset C \\ & \Rightarrow y = \lambda(x_1 + \frac{1}{\lambda}v) + (1 - \lambda)x_2 \in C \end{aligned}$$

by convexity.

Next, assume that $x_2 \in \bar{C}$*. Then there exists some* $\bar{x}_2 \in C$ *such that* $||\bar{x}_2 - x_2|| < \frac{\lambda r}{1-\lambda}$

Let

$$v = x_1 - \left(\frac{1-\lambda}{\lambda}\right)(\bar{x}_2 - x_2)$$

Then

$$\begin{aligned} ||v - x_1|| & = \frac{1-\lambda}{\lambda}||\bar{x}_2 - x_2|| \\ & \leq \frac{1-\lambda}{\lambda}\frac{\lambda r}{1-\lambda} \\ & = r \\ & \Rightarrow v \in B(x_1, r) \subset C \end{aligned}$$

But

$$\begin{aligned} & \lambda v + (1-\lambda)\bar{x}_2 \\ = & \ \lambda\left(x_1 - \left(\frac{1-\lambda}{\lambda}\right)(\bar{x}_2 - x_2)\right) \\ & +(1-\lambda)\bar{x}_2 \\ = & \ \lambda x_1 + (1-\lambda)x_2 = z \end{aligned}$$

thus $z \in C^o$ ■

Two immediate corollaries of this are

Corollary 1 *If C is convex in a linear space, then so is C^o*

Corollary 2 *If C is convex in a linear space and $C^o \neq \{\}$, then*

1. $cl(C^o) = cl(C)$
2. $C^o = (cl(C))^o$

Notice that this second is not true for any arbitrary set A. For example, let

$$A = [\mathbb{Q} \cap [0,1]] \cup [1,2]$$

then $int(A) = (1,2)$ and $(cl(A))^o = (0,2)$

So we know that if a set is convex, then so is its interior. It turns out that its closure will also be convex

Lemma 3 *If C is convex, then so is $\bar{C}$.*

Proof. *Let $x, y \in \bar{C}$ and $\lambda \in (0,1)$. We need to show that $z = \lambda x + (1-\lambda)y \in \bar{C}$. Let $\{x_k\}$ and $\{y_k\}$ be sequences in $\bar{C}$ that converge to x and y respectively. We know that*

$$\lambda x_k + (1-\lambda)y_k \in C \;\forall\; k$$

by convexity, and that $\lambda x_k + (1-\lambda)y_k \to z$. Thus, as the closure is closed, $z \in \bar{C}$. ■

We will finish off this section by showing that two properties translate from a set to its convex hull.

Lemma 4 *If O is open in some linear space, then so is $co(O)$*

Proof. *Assume that $O \neq \{\}$ (otherwise it is trivial). Then $O \subseteq co(O) \Rightarrow O = O^o \subseteq [co(O)]^o \neq \{\}$*

So the interior of the convex hull is a convex set that contains O, thus, as the convex hull in the intersection of such sets,

$$co(O) \subset int(co(O))$$

but as int(co(O)) ⊂ co(O) trivially, we have co(O) = int(co(O)), and so co(O) is open. ■

Lemma 5 *If $K \subset \mathbb{R}^n$ is compact, then so is $co(K)$*

Proof. *Let $\theta : \mathbb{R}^{n+1} \times (\mathbb{R}^n)^{n+1} \to \mathbb{R}^n$ be defined as*

$$\theta(\lambda, x_1, ..., x_{n+1}) = \sum_{i=1}^{n+1} \lambda_i x_i$$

You can check (but it should be obvious) that θ is continuous. Now note that

$$\begin{aligned} & co(k) \\ = & \left\{ \sum_{i=1}^{n+1} \lambda_i x_i | \lambda \in P_{n+1}, \{x_1, x_2, ..., x_{n+1}\} \in K^{n+1} \right\} \end{aligned}$$

by Carathedoery, so

$$co(k) = \theta(P_{n+1,} K^{n+1})$$

both of which are compact sets (as P_{n+1}) is a closed, bounded subset of $\mathbb{R}^{n+1}$, and K^{n+1} is the product of compact sets. Thus, $co(K)$ is the the image of a continuous function on a compact pre-image and so, (as we have proved) is compact itself. ■

2.2 Orthogonal Projection and Convexity

We are now going to come back to the concept of orthogonal projection we discussed in linear algebra. There we showed that, for any linear subspace S and point x, there was a unique decomposition $s_x \in S$ and $s_x^\perp \in S^\perp$ such that $s_x + s_x^\perp = x$. Moreover, we said that s_x was the closest point to x in the subspace S. We are now going to use this latter notion to define orthogonal projection more generally.

Definition 4 *Let $C \subset \mathbb{R}^n$ and $y \in \mathbb{R}^n$. Then if $x^* \in C$ is such that $||y - x^*|| \leq ||y - x|| \; \forall \; x \in C$, then x^* is the* ***orthogonal projection*** *of y onto C, and we write $x^* = P_c(y)$*

We are now going to prove a result which is similar to Hilbert's projection theorem (we are going to prove it for $\mathbb{R}^n$, but the result is more general). This will show that, for closed convex sets, the orthogonal projection exists and is unique. This section is going to rely heavily on the result we proved in the linear analysis section:

Lemma 6 *For* $x, y \in \mathbb{R}^n$, $||x+y||^2 = ||x||^2 + ||y||^2 + 2 < x, y >$

Proof. *See linear algebra lecture notes* ■

We will begin by showing that, for any convex set, there can be at most one element in the convex projection.

Lemma 7 *If* C *is convex, then there exists at most one* $x^* \in C$ *such that* $||y - x^*|| \leq ||y - x|| \; \forall$ $x \in C$

Proof. Proof. Say that two vectors, x_1 and x_2 both satisfy this condition. First note, that for any two vectors x_1, x_2, y we already know that

$$||x_1 - x_2||^2 = ||x_1 - y||^2 + ||x_2 - y||^2 - 2 < (x_1 - y).(x_2 - y) >$$

But also

$$\begin{aligned} 4||\frac{x_1 + x_2}{2} - y||^2 &= ||(x_1 - y) + (x_2 - y)|| \\ &= ||x_1 - y||^2 + ||x_2 - y||^2 + 2 < (x_1 - y).(x_2 - y) > \end{aligned}$$

And so

$$||x_1 - x_2||^2 = 2||x_1 - y||^2 + 2||x_2 - y||^2 - 4||\frac{x_1 + x_2}{2} - y||^2$$

Let δ be the distance between x_1 and y. as $\frac{x_1+x_2}{2} \in C$, we know that $||\frac{x_1+x_2}{2} - y|| \geq \delta$, and so

$$\begin{aligned} & ||x_1 - x_2||^2 \\ \leq \; & 2\delta^2 + 2\delta^2 - 4\delta^2 \\ \Rightarrow \; & ||x_1 - x_2||^2 = 0 \end{aligned}$$

Thus $x_1 = x_2$ ■ ■

So convexity is enough to imply uniqueness of an orthogonal projection. Is it enough to guarantee existence? Of course not! Consider the convex set $(0, 1)$ and the point $y = 2$. However, in $\mathbb{R}^n$, it turns out that closedness is enough to guarantee existence.

Theorem 3 *Let $C \subset \mathbb{R}^n$ be convex and closed, and $y \in \mathbb{R}^n$. Then $P_c(y)$ exists, is unique and $x^* = P_c(y)$ iff $x^* \in C$ and*

$$< y - x^*, x - x^* > \leq 0 \ \forall \ x \in C$$

Proof. *First we will prove existence. Let $m = \inf_{x \in C} ||y - x||$ and $R = 2m$ and $C_R = C \cap \bar{B}(y, R)$. As this is an intersection between two convex sets it is convex. As it is the intersection of two closed sets is is closed, and as it is also bounded it is compact. It is also non-empty. As $f(x) = ||y - x||$ is continuous in $\mathbb{R}^n$, then by Weierstrass, $m = \inf_{x \in C_R} f(x)$ is attained in C_R. let $x^* = P_{C_R}(y) = P_C(y)$. Thus we have existence and (by lemma 7) uniqueness*

Now we need to prove (first) that if $x^* = P_C(y)$, then $< y - x^*, x - x^* > \leq 0 \ \forall \ x \in C$. To see this, pick an $x \in C$ and $\lambda \in [0, 1]$. We know that

$$\begin{aligned} ||x^* - y||^2 &\leq ||y - (\lambda x + (1 - \lambda)x^*)||^2 \\ &= ||y - x^* - \lambda(x - x^*)||^2 \\ &= ||y - x^*||^2 + \lambda^2 ||(x - x)^*||^2 - 2\lambda < y - x^*, x - x^* > \\ &\Rightarrow < y - x^*, x - x^* > \leq \frac{\lambda}{2} ||(x - x)^*||^2 \end{aligned}$$

This is true for all λ, thus taking the limit as λ goes to zero we can see

$$< y - x^*, x - x^* > \leq 0$$

Finally, we need to show that if $< y - x^*, x - x^* > \leq 0 \ \forall \ x \in C$, then $x^* = P_C(y)$. To see this, note that, for any $x \in C$,

$$\begin{aligned} ||y - x||^2 &= ||(y - x^*) - (x - x^*)||^2 \\ &= ||y - x^*||^2 + ||x - x^*||^2 - 2 < y - x^*, x - x^* > \end{aligned}$$

but, as $||x - x^*||^2 \geq 0$ and $-2 < y - x^*, x - x^* > \geq 0$

$$||y - x|| \geq ||y - x^*||$$

■