Math 591 – Real Algebraic Geometry and Convex Optimization

Lecture 1: Convexity basics

Cynthia Vinzant, Spring 2019

We take V to be a vector space over $\mathbb{R}$. For example $\mathbb{R}^n$, the polynomial ring $\mathbb{R}[x_1, \ldots, x_n]$, or the vector space of real valued continuous function $f : [0,1] \to \mathbb{R}$. Mostly we will deal with $\dim(V) < \infty$, but it's good to keep the other examples in mind.

Definition. A set $C \subseteq V$ is **convex** if for any $u, v \in C$, C contains the line segment between u and v. That is, for any $\lambda \in [0,1]$, $\lambda u + (1-\lambda)v \in C$.

Example.

- the half-open disk $\{(x,y) : x^2 + y^2 < 1\} \cup \{(x,y) : x^2 + y^2 = 1, y \leq 0\}$
- $\{$continuous $f : [0,1] \to \mathbb{R}$ such that $f(r) \geq 0$ for all $r \in [0,1]$ and $f(1) = 1\}$

A **convex combination** of points $v_1, \ldots, v_k \in V$ is a point of the form $\sum_{i=1}^k v_i$ where $\lambda_i \geq 0$ and $\sum_{i=1}^k \lambda_i = 1$, and the **convex hull** of a subset $S \subseteq V$ is defined as the set of all convex combinations of finite subsets of V:

$$\text{conv}(S) = \left\{ \sum_{i=1}^k \lambda_i v_i : k \in \mathbb{N}, v_1, \ldots, v_k \in S, \text{ and } \lambda_1, \ldots, \lambda_k \geq 0 \text{ with } \sum_{i=1}^k \lambda_i = 1 \right\}.$$

Proposition. *For any $S \subseteq V$, $\text{conv}(S)$ is convex.*

Proof. Let $\lambda \in [0,1]$, and to take two arbitrary elements of $\text{conv}(S)$, we take two collections of points $v_1, \ldots, v_k$, $w_1, \ldots, w_\ell \in S$ and nonnegative scalars $\alpha_1, \ldots, \alpha_k$, $\beta_1, \ldots, \beta_\ell$ with $\sum_{i=1}^k \alpha_i = 1$ and $\sum_{j=1}^\ell \beta_j = 1$. Then we see that

$$\lambda \sum_{i=1}^k \alpha_i v_i + (1-\lambda) \sum_{j=1}^\ell \beta_j w_j \;=\; \sum_{i=1}^k \lambda \alpha_i v_i + \sum_{j=1}^\ell (1-\lambda)\beta_j w_j.$$

Since each coefficient $\lambda\alpha_i$ and $(1-\lambda)\beta_j$ is nonnegative and they all sum to one,

$$\sum_{i=1}^k \lambda \alpha_i + \sum_{j=1}^\ell (1-\lambda)\beta_j \;=\; \lambda \sum_{i=1}^k \alpha_i + (1-\lambda) \sum_{j=1}^\ell \beta_j \;=\; \lambda + (1-\lambda) = 1,$$

this is a convex combination of finitely many points from S and thus in $\text{conv}(S)$. □

If S is finite, then $\text{conv}(S)$ is called a **polytope**.

Definition. An **affine combination** of points $v_1, \ldots, v_k \in V$ is a point of the form $\sum_{i=1}^k v_i$ where $\lambda_i \in \mathbb{R}$ and $\sum_{i=1}^k \lambda_i = 1$, and the **affine hull** of a subset $S \subseteq V$ is defined as the set of all affine combinations of finite subsets of V:

$$\text{aff}(S) = \left\{ \sum_{i=1}^k \lambda_i v_i : k \in \mathbb{N}, v_1, \ldots, v_k \in S, \text{ and } \lambda_1, \ldots, \lambda_k \in \mathbb{R} \text{ with } \sum_{i=1}^k \lambda_i = 1 \right\}.$$

We define the dimension of a convex set C to be the dimension of its affine hull, $\text{aff}(C)$, as an affine linear space.

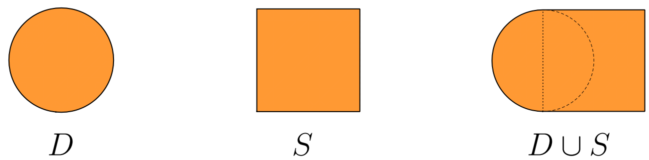
Definition. A subset $F \subseteq C$ of a convex set C is a **face** of C if F is convex and has the property that

$$\lambda u + (1-\lambda)v \in F \quad \Rightarrow \quad u, v \in F$$

for all $\lambda \in (0,1)$ and $u, v \in C$. Both $\emptyset$ and C are always faces of C. All other faces we call **proper** faces of C.

Question. What are the proper faces of the three convex sets D, S, and $D \cup S$ where

$$D = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}, \text{ and}$$
$$S = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 2, -1 \leq y \leq -1\}?$$



For those familiar with polytopes, faces are often defined by the sets maximizes a linear function. Recall that the dual vector space of V, V^*, is the real vector space of linear functionals $\ell : V \to \mathbb{R}$.

Proposition. *For any* $\ell \in V^*$ *and convex set* $C \subseteq V$*, the set*

$$F = \{v \in C : \ell(v) \geq \ell(w) \text{ for all } w \in C\}$$

is a face of C*.*

In this case we say that ℓ **exposes** F and that F is an **exposed face** of C.

Example. In both the disk and the square from above, all faces are exposed. For example, in the disk D the face $\{(1,0)\}$ is exposed by the linear function $\ell(x,y) = y$. This function also exposes the face $[0,2] \times \{1\}$ in the square S, where as the face $\{(1,2)\}$ is exposed by $\ell(x,y) = x + y$.

Non-example. Now consider the face $\{(0,1)\}$ of the union $D \cup S$. The only linear functions achieving their maximum over C at $(0,1)$ have the form $\ell(x,y) = cy$ for $c \geq 0$. But any such function exposes the face $[0,2] \times \{1\}$.

One question we address next time will be: given a point $v \in C$, what linear functions attain their maximum (over C) at v?