10-425/625: Introduction to Convex Optimization (Fall 2023)

Lecture 3: Convex Sets

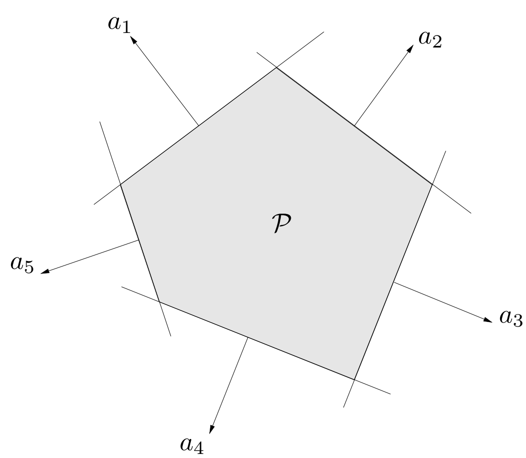
Instructor:[1] *Matt Gormley* *September 6, 2023*

3.1 Convex Sets (continued)

3.1.1 Examples of Convex Sets

Some more examples (again, useful to make sure you know how to verify the convexity of these sets):

8. **Polyhedra:** The set $\{x : Ax \leq b\}$ for given A, b (or equivalently, sets of the form $\{x : Ax \leq b, Cx = d\}$).



Background:

Definition 3.1 (Linear independence)**.** *Vectors* $\{x_1, \ldots, x_k\}$ *are* linearly independent *if there is no* $\lambda_1, \ldots, \lambda_k$ *such that* $\sum_{i=1}^{k} \lambda_i x_i = 0$ *except all zeros.*

[1] These notes were originally written by Siva Balakrishnan for 10-725 Spring 2023 (original version: here) and were edited and adapted for 10-425/625.

> **Definition 3.2** (Affine independence)**.** *Vectors* $\{x_1, \ldots, x_k\}$ *are* affinely independent *if there is no* $\lambda_1, \ldots, \lambda_k$*, with* $\sum_{i=1}^k \lambda_i = 0$ *such that* $\sum_{i=1}^k \lambda_i x_i = 0$ *except all zeros.*

9. **Simplices:** For a collection of affinely independent points $x_1, \ldots, x_k$, the corresponding simplex is simply the convex hull $\text{conv}\{x_1, \ldots, x_k\}$.

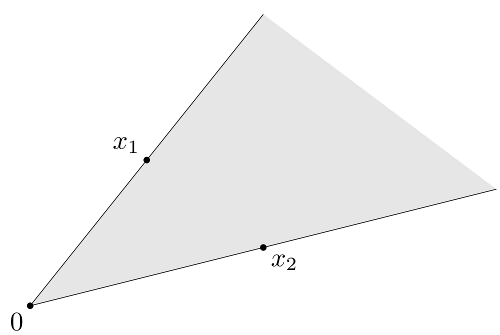
   A prominent example is the probability simplex, which is the convex hull of the d basis vectors $e_1, \ldots, e_d$.

3.1.2 Convex Cones

> **Background: (Positive Definite and Positive Semidefinite)**
> Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix.
>
> **Definition 3.3.** *The matrix A is* positive semidefinite*, written* $A \succeq 0$*, if* $x^T A x \geq 0$ *for all* $x \in \mathbb{R}^n$*.*
>
> **Definition 3.4.** *The matrix A is* positive definite*, written* $A \succ 0$*, if* $x^T A x > 0$ *for all non-zero* $x \in \mathbb{R}^n$*.*

A set C is a *cone* if for every $x \in C$, $\theta x \in C$ for any $\theta \geq 0$, i.e. for any point in C the ray joining that point to the origin must also be in C. Cones are not convex in general, so we will refer to *convex cones* as cones which are additionally convex.

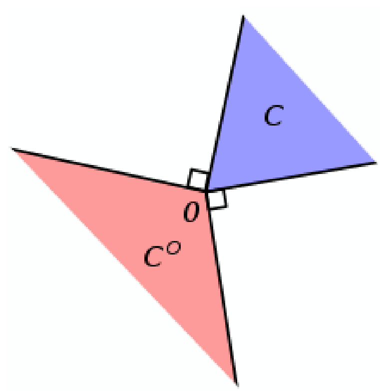
In the example below, the two rays are a cone; whereas the shaded region is a convex cone.



It is easy to see that convex cones additionally satisfy the property that if $x_1, x_2 \in C$ then for any $\theta_1, \theta_2 \geq 0$, $\theta_1 x_1 + \theta_2 x_2 \in C$. These are called *conic* *combinations*, i.e. for $x_1, \ldots, x_k$, a conic combination is any point of the form $\theta_1 x_1 + \ldots + \theta_k x_k$ with $\theta_i \geq 0$ is called a conic combination. The conic hull of a set C collects all conic combinations of points in C, and is the smallest *convex* cone containing C.

There are several important cones:

1. **Norm Cone:** $\{(x,t) : \|x\| \leq t\}$. For the ℓ_2 norm this cone is called the second-order cone (sometimes called the ice-cream cone).
2. **PSD Cone:** Denoted $\mathbb{S}^d_+ = \{X \in \mathbb{S}^d : X \succeq 0\}$, i.e. X is a symmetric matrix, with all positive eigenvalues.
3. **Polar Cone:** For any cone C, the *polar* cone C° is defined as the collection of vectors which make an atleast 90-degree angle with all vectors in C, i.e.
$$C^\circ = \{x : x^T y \leq 0, \text{for all } y \in C\}.$$



There is a fundamental reason why cones will be important to us. We will use them to characterize optimality. Two cones are important in this context: the normal cone and its polar cone (which has its own name, the tangent cone).

4. **Normal Cone:** *(definition saved for a later time)*
5. **Polar Cone:** *(definition saved for a later time)*

3.2 The Separating and Supporting Hyperplane Theorems

Background: (Open and Closed Sets)

Consider a set $S \subseteq \mathbb{R}^n$.

Definition 3.5 (Interior Point)**.** *For a set $S \subseteq \mathbb{R}^n$, an element of that set $x \in S$ is an* interior point *if there exists an epsilon-ball around x that is entirely within the set S:*

$$\{y : ||y - x||_2 \leq \epsilon\} \subseteq C$$

Definition 3.6 (Open Set and Closed Set)**.** *A set $S \subseteq \mathbb{R}^n$ is* open *if all points in S are interior points. A set $S \subseteq \mathbb{R}^n$ is* closed *if its complement $C^c = \{x \in \mathbb{R}^n : x \notin C\}$ is open.*

For example, the interval $(0, 1)$ on the real line is an open set, whearas the interval $[0, 1]$ is a closed set. In 2D, the set of points $x \in \mathbb{R}^2$ satisfying $x_1 + x_2 > 7$ is an open set, whereas the set of points satisfying $x_1 + x_2 \geq 7$ is a closed set.

Definition 3.7 (Boundary Point)**.** *We say that a* boundary point *satisfies the property that are points both in S and not in S that are arbitrarily close. That is, $x \in \mathbb{R}^n$ is a boundary point of S if for all $\epsilon > 0$, $\exists y \in S$ and $\exists z \notin S$ such that*

$$\begin{aligned} &||y - x||_2 \leq \epsilon \textit{ and} \\ &||z - x||_2 \leq \epsilon \end{aligned}$$

Definition 3.8 (Boundary)**.** *The* boundary *of a set $S \subseteq \mathbb{R}^n$ are all points in $\mathbb{R}^n$ that are boundary points.*

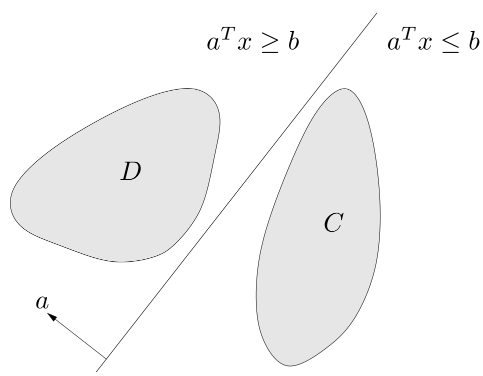
The above definition of a boundary has two consequences of note: First, all points in S that are not interior points are boundary points. Second, not all boundary points are in S, e.g. an open set contains none of its boundary points.

We can characterize a closed set in two additional ways:

1. A set $S \subseteq \mathbb{R}^n$ is closed if it contains all its boundary points.
2. A set $S \subseteq \mathbb{R}^n$ is closed if for every limiting sequence of points $x_1, x_2, x_3, \ldots$ that converges to x, $x_i \in S \Rightarrow x \in S$.

Theorem 3.9 (Separating Hyperplane)**.** *If C and D are non-empty convex sets which are disjoint, i.e. $C \cap D = \emptyset$, then there exists a separating hyperplane, i.e. a, b such that,*

$$a^T x \leq b, \ \text{for all } x \in C,$$
$$a^T x \geq b, \ \text{for all } x \in D.$$

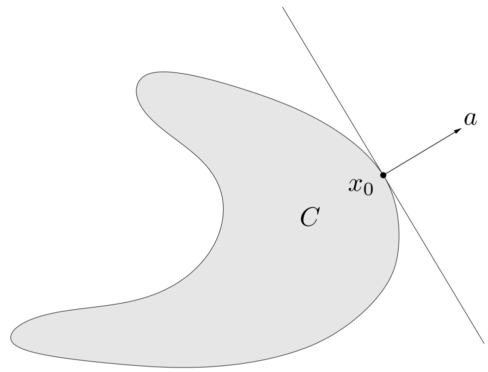


Notice that, it is *not* generally true of two disjoint nonconvex sets that there exists a separating hyperplane.

Theorem 3.10 (Supporting Hyperplane)**.** *If C is a non-empty convex set, and $x_0 \in boundary(C)$, then there is a vector a such that,*

$$a^T (x - x_0) \leq 0, \ \text{for all } x \in C.$$

The latter has an interesting converse, if the set C is closed (check what this means if you're not familiar with it), and has a non-empty interior, and has a supporting hyperplane at every point then C must be convex.



The proofs of these theorems (at least in the case where the sets are closed and bounded) is straightforward (and explicit) – see BV, Section 2.5 if you are curious.

Segue... Next time, we'll talk operations that preserve convexity of a set and begin our discussion of convex functions.