Lecture 7

"[A]lmost every "convex" idea can be explained by a two-dimensional picture."

— Alexander Barvinok

In this lecture we begin our study of the theory underlying constrained convex optimization. One way to define a *convex optimization problem* is

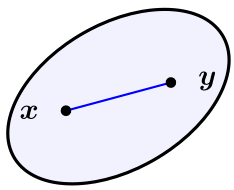
$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & g_1(\boldsymbol{x}) \leq 0 \\ & \cdots \\ & g_m(\boldsymbol{x}) \leq 0 \\ & \boldsymbol{x} \in \Omega \end{aligned}$$

where $f, g_1, \ldots, g_m \colon \mathbb{R}^d \to \mathbb{R}$ are *convex* functions and $\Omega \subseteq \mathbb{R}^d$ is a *convex* set. The special case where the f and the g_i are linear functions and $\Omega = \mathbb{R}^d$ is known as linear programming, and is studied first. Before embarking on the study of models and algorithms for convex optimization, we need to study convex sets in more depth.

7.1 Convex sets

We recall the definition of a convex set.

Definition 7.1. A set $C \subseteq \mathbb{R}^d$ is a *convex set*, if for all $\boldsymbol{x}, \boldsymbol{y} \in C$ and $\lambda \in [0, 1]$, $\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} \in C$. In words, for every two points in C, the line joining them is also in C. A compact (closed and bounded) convex set is called a *convex body*.



We will denote by $\mathcal{C}(\mathbb{R}^d)$ the collection of convex sets and by $\mathcal{K}(\mathbb{R}^d)$ the collection of convex bodies. The following Lemma is left as an exercise.

Lemma 7.2. *Let $C, D \in \mathcal{C}(\mathbb{R}^d)$ be convex sets. Then the following are also convex.*

- $C \cap D$;
- $C + D = \{\boldsymbol{x} + \boldsymbol{y} : \boldsymbol{x} \in C, \boldsymbol{y} \in D\}$;
- $\boldsymbol{A}C = \{\boldsymbol{A}\boldsymbol{x} : \boldsymbol{x} \in C\}$, *where* $\boldsymbol{A} \in \mathbb{R}^{m\times d}$.

The *convex hull* $\operatorname{conv} S$ of a set S is the intersection of all convex sets containing S. Clearly, if S is convex, then $S = \operatorname{conv} S$.

Example 7.3. Let $S = \{(1,1)^\top, (1,-1)^\top, (-1,1)^\top, (-1,-1)^\top, (0,0)^\top\}$. The convex hull of this set is the square.

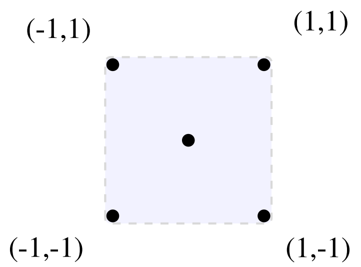


Figure 7.1: A convex hull of five points.

A *convex combination* of points $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k$ is a linear combination

$$\sum_{i=1}^k \lambda_i \boldsymbol{x}_i$$

such that $\lambda_i \geq 0$ and $\sum_{i=1}^k \lambda_i = 1$. It can be shown inductively that convex sets are closed under convex combinations: any convex combination of points in $C \in \mathcal{C}(\mathbb{R}^d)$ is still in C. In fact, the set of all convex combinations of points in a set S is the convex hull of S.

Lemma 7.4. *Let S be a set. Then*

$$\operatorname{conv} S = \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{x} = \sum_{i=1}^k \lambda_i \boldsymbol{x}_k,\ \boldsymbol{x}_i \in S,\ \sum_{i=1}^k \lambda_i = 1,\ \lambda_i \geq 0\}.$$

Example 7.5. A hyperplane, defined as the solution set of one linear equations,

$$H = \{\boldsymbol{x} : \langle \boldsymbol{a}, \boldsymbol{x} \rangle = b\},$$

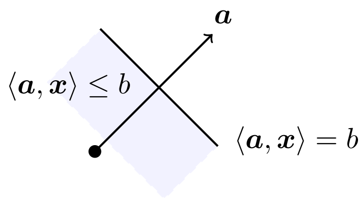


Figure 7.2: Hyperplane and halfspace

is a convex set. Define the halfspaces H_+ and H_- as the two sides that H divides $\mathbb{R}^d$ into:

$$H_- = \{\boldsymbol{x} : \langle \boldsymbol{a}, \boldsymbol{x} \rangle \leq b\}, \quad H_+ = \{\boldsymbol{x} : \langle \boldsymbol{a}, \boldsymbol{x} \rangle \geq b\}$$

These are also convex sets.

Example 7.6. Euclidean balls and ellipsoids are common examples of convex sets. Let $\boldsymbol{P}$ be a positive semidefinite symmetric matrix. Then an ellipsoid with center $\boldsymbol{x}_0$ is a set of the form

$$\mathcal{E} = \{\boldsymbol{x} : \langle \boldsymbol{x} - \boldsymbol{x}_0, \boldsymbol{P}^{-1}(\boldsymbol{x} - \boldsymbol{x}_0) \rangle \leq 1\}.$$

A Euclidean unit ball is the special case $\boldsymbol{P} = \boldsymbol{I}$.

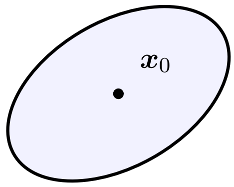


Figure 7.3: An ellipse

Example 7.7. A *convex cone* is a set C such that for all $\boldsymbol{x}, \boldsymbol{y}$ and $\lambda \geq 0, \mu \geq 0$, $\lambda \boldsymbol{x} + \mu \boldsymbol{y} \in C$. It is easily verified that such a set is convex. Three important cones are the following:

1. The non-negative orthant $\mathbb{R}^d_+ = \{\boldsymbol{x} \in \mathbb{R}^d : x_i \geq 0, 1 \leq i \leq d\}$,
2. The second order (ice cream) cone (or Lorentz cone)

$$C_\alpha = \{\boldsymbol{x} : \sum_{i=1}^{d-1} x_i^2 \leq x_d^2\},$$

3. The cone $\mathcal{S}^d_+$ of positive semidefinite symmetric matrices.

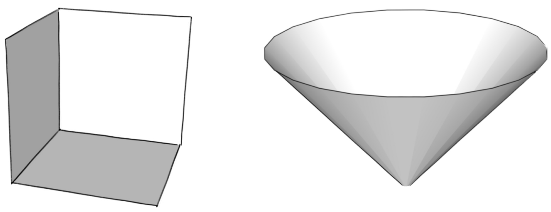


Figure 7.4: The orthant and the second order cone

Possibly the most important result in convex geometry is the *hyperplane separation theorem*. We first need the following.

Lemma 7.8. *Let C be a non-empty convex set and $\boldsymbol{x} \notin C$. Then there exists a point $\boldsymbol{y} \in C$ that minimizes the distance $\|\boldsymbol{x} - \boldsymbol{y}\|$. Moreover, for all $\boldsymbol{z} \in C$ we have*

$$\langle \boldsymbol{z} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle \leq 0.$$

In words, the vectors $\boldsymbol{z} - \boldsymbol{y}$ and $\boldsymbol{x} - \boldsymbol{y}$ form an obtuse angle.

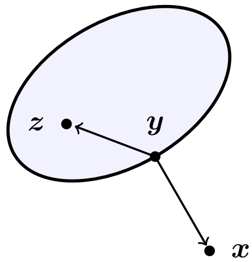


Figure 7.5: Internal and external directions

Proof. Since $C \neq \emptyset$, there exists $r > 0$ such that the ball $B(\boldsymbol{x}, r) := \{\boldsymbol{y} \in \mathbb{R}^d : \|\boldsymbol{y} - \boldsymbol{x}\| \leq \varepsilon\}$ intersected with C is not empty. Since $K := C \cap B(\boldsymbol{x}, r)$ is compact (closed and bounded) and the function $\|\boldsymbol{y} - \boldsymbol{x}\|$ is continuous on K, it has a minimizer $\boldsymbol{y} \in K$. For the second claim, note that since C is convex, for every $\lambda \in [0, 1]$,

$$\boldsymbol{w} = \lambda \boldsymbol{z} + (1 - \lambda)\boldsymbol{y} \in C.$$

For the distance between $\boldsymbol{z}$ and $\boldsymbol{x}$ we then get

$$\begin{aligned}
\|\boldsymbol{w} - \boldsymbol{x}\|^2 &= \|\lambda \boldsymbol{z} + (1 - \lambda)\boldsymbol{y} - \boldsymbol{x}\|^2 = \|\lambda(\boldsymbol{z} - \boldsymbol{y}) - (\boldsymbol{x} - \boldsymbol{y})\|^2 \\
&= \lambda^2 \|\boldsymbol{z} - \boldsymbol{y}\|^2 - 2\lambda \langle \boldsymbol{z} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle + \|\boldsymbol{x} - \boldsymbol{y}\|^2.
\end{aligned}$$

We now prove the claim by contradition. Assume $\langle \boldsymbol{z} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle > 0$. Then we can choose λ such that

$$0 < \lambda < \min \left\{ \frac{2\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{z} - \boldsymbol{y} \rangle}{\|\boldsymbol{z} - \boldsymbol{y}\|^2}, 1 \right\}.$$

With such a λ we get

$$\|\boldsymbol{w}-\boldsymbol{x}\|^2 = \lambda^2\|\boldsymbol{z}-\boldsymbol{y}\|^2 - 2\lambda\langle \boldsymbol{z}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y}\rangle + \|\boldsymbol{x}-\boldsymbol{y}\|^2 < \|\boldsymbol{x}-\boldsymbol{y}\|^2.$$

This inequality, however, contradicts the assumption that $\boldsymbol{y}$ is a closest point, so that $\langle \boldsymbol{z}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y}\rangle \leq 0$ has to hold. $\square$

In what follows write $\mathrm{int}S$ for the *interior* of a set S.

Theorem 7.9. *Let C be a closed convex set and $\boldsymbol{x} \notin C$. Then there exists a hyperplane H such that $C \subset \mathrm{int}H_-$ and $\boldsymbol{x} \in \mathrm{int}H_+$.*

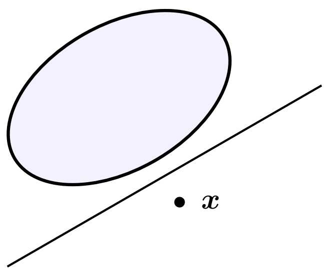


Figure 7.6: A separating hyperplane

Proof. Let $\boldsymbol{y} \in C$ be a nearest point to $\boldsymbol{x}$ in C, i.e., a point such that for all other $\boldsymbol{z} \in C$, $\|\boldsymbol{x}-\boldsymbol{y}\| \leq \|\boldsymbol{x}-\boldsymbol{z}\|$. Define

$$\boldsymbol{a} = \boldsymbol{x}-\boldsymbol{y}, \quad b = (\|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2)/2.$$

We aim to show that $\langle \boldsymbol{a}, \boldsymbol{x}\rangle = b$ defines a separating hyperplane.

For this we have to show that

1. $\langle \boldsymbol{a}, \boldsymbol{x}\rangle > b$;
2. For all $\boldsymbol{z} \in C$, $\langle \boldsymbol{a}, \boldsymbol{z}\rangle < b$.

For (1), note that

$$\langle \boldsymbol{a}, \boldsymbol{x}\rangle = \langle \boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}\rangle > \langle \boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}\rangle - \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{y}\|^2 = \frac{1}{2}(\|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2) = b.$$

To prove (2), assume on the contrary that there exists a $\boldsymbol{z} \in C$ such that $\langle \boldsymbol{a}, \boldsymbol{z}\rangle \geq b$. We know that the point $\boldsymbol{y} \in C$ satisfies the inequality (2), since

$$\langle \boldsymbol{a}, \boldsymbol{y}\rangle < \langle \boldsymbol{a}, \boldsymbol{y}\rangle + \frac{1}{2}\|\boldsymbol{a}\|^2 = \langle \boldsymbol{a}, \boldsymbol{y}\rangle + \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{y}\|^2 = b.$$

Therefore,

$$\langle \boldsymbol{a}, \boldsymbol{z}-\boldsymbol{y}\rangle = \langle \boldsymbol{a}, \boldsymbol{z}\rangle - \langle \boldsymbol{a}, \boldsymbol{y}\rangle > b - b = 0,$$

but this contradict Lemma 7.8. We therefore conclude $\langle \boldsymbol{a}, \boldsymbol{z}\rangle < b$. The separating hyperplane H is thus defined by the equation $\langle \boldsymbol{a}, \boldsymbol{x}\rangle = b$. $\square$