1.1 *Lecture 1. Affine Sets and Convex Sets.*

1.1.1 Overview

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable and consider the following problem:

$$\begin{array}{rl} \min & f(x) \\ s.t. & x \in C \subseteq \mathbb{R}^n \end{array} \qquad (P)$$

We call f the **objective function** and C the **constraint set**.

In the special case where $C = \mathbb{R}^n$, the minimizers of f (if any) will occur at the *critical points* of f, namely the points $x \in \mathbb{R}^n$ such that $\nabla f(x) = 0$. This is known as the **Fermat's rule**, which we will learn about more later.

In this course, we will discuss and learn convexity of sets and functions and how we can approach problem (P) in the more general settings of:

1. absence of differentiability of the function f, where f is convex, and/or
2. $\varnothing \neq C \subsetneq \mathbb{R}^n$, where C is convex.

1.1.2 Affine Sets and Affine Subspaces in $\mathbb{R}^n$

Geometrically speaking, a non-empty subset $S \subseteq \mathbb{R}^n$ is **affine** if the *line* [1] connecting any two points in the set lies entirely in the set. The intuitive picture is that of an endless uncurved structure, like a line or a plane in space. [2]

Definition 1.1. Let $S \subseteq \mathbb{R}^n$. Then

- S is an **affine set** if $\lambda x + (1 - \lambda)y \in S$ for all $x, y \in S$ and $\lambda \in \mathbb{R}$.
- S is an **affine subspace** if it is a *non-empty* affine set.
- Let $S \subseteq \mathbb{R}^n$. Then **affine hull** of S, denoted by $\operatorname{aff}(S)$, is the intersection of all affine sets containing S. Equivalently, it is the smallest affine set containing S.

Some elementary examples of affine sets of $\mathbb{R}^n$:

1. L, where $L \subseteq \mathbb{R}^n$ is a linear subspace.
2. $a + L$, where $a \in \mathbb{R}^n$ and $L \subseteq \mathbb{R}^n$ is a linear subspace.
3. $\varnothing$ and $\mathbb{R}^n$ are extreme affine sets of $\mathbb{R}^n$.

[1] Not line segment! That is for convex sets.

[2] Another perspective: An affine space is what is left of a vector space after you've forgotten which point is the origin. Marcel Berger: "An affine space is nothing more than a vector space whose origin we try to forget about, by adding translations to the linear maps." Connect this with examples 1 and 2.

1.1.3 Convex Sets in $\mathbb{R}^n$

A subset $C \subseteq \mathbb{R}^n$ is **convex** if given any two points x and y in C, the *line segment* joining x and y lies entirely in C.

Definition 1.2. A subset C of $\mathbb{R}^n$ is **convex** if

$$\lambda x + (1 - \lambda) y \in C$$

for all $x, y \in C$ and $0 < \lambda < 1$.

Theorem 1.3. *The intersection of an arbitrary collection of convex sets is convex.*

Proof. Let $(C_i)_{i \in I}$ be a collection of convex subsets of $\mathbb{R}^n$ indexed by I. Define $C := \bigcap_{i \in I} C_i$. Fix $x, y \in C$ and $\lambda \in (0, 1)$. Since C_i is convex for all $i \in I$, i.e.,

$$\forall i \in I : \lambda x + (1 - \lambda) y \in C_i,$$

we get $\lambda x + (1 - \lambda) y \in \bigcap_{i \in I} C_i = C$. Hence, C is convex. □

Corollary 1.4. *Let $b_i \in \mathbb{R}^n$ and $\beta_i \in \mathbb{R}$ for $i \in I$, where I is an arbitrary index set. Then the set*

$$C = \{x \in \mathbb{R}^n \mid \forall i \in I : \langle x, b_i \rangle \leq \beta_i\}$$

is convex.

Proof. For each $i \in I$, define

$$C_i := \{x \in \mathbb{R}^n \mid \langle x, b_i \rangle \leq \beta_i\}.$$

We claim that all such C_i's are convex. Indeed, let $i \in I$ and fix $x, y \in C_i$ and $\lambda \in (0, 1)$. Set $z := \lambda x + (1 - \lambda) y$. Then

$$\begin{aligned}
\langle z, b_i \rangle &= \langle \lambda x + (1 - \lambda y), b_i \rangle \\
&= \lambda \langle x, b_i \rangle + (1 - \lambda) \langle y, b_i \rangle && \text{linearity of } \langle \cdot, \cdot \rangle \\
&\leq \lambda \beta_i + (1 - \lambda) \beta_i && \forall x \in C_i : \langle x, b_i \rangle \leq \beta_i \\
&= \beta_i.
\end{aligned}$$

Thus, $z \in C_i$ and C_i is convex. The result follows from Theorem 1.3. □

1.1.4 Convex Combination of Vectors

Definition 1.5. A vector sum $\lambda_1 x_1 + \cdots + \lambda_m x_m$ is called a **convex combination** of vectors $x_1, \ldots, x_m$ if

$$\forall i \in [m] : \lambda_i \geq 0 \quad \text{and} \quad \sum_{i=1}^{m} \lambda_i = 1.$$

Theorem 1.6. *A subset $C \subseteq \mathbb{R}^n$ is convex iff it contains all the convex combinations of its elements.*

Proof. $\Leftarrow$: Trivial. $\Rightarrow$: Suppose C is convex. We proceed by induction on m, the number of vectors in the convex combination. For $m = 2$, the conclusion is clear as C is convex. Now suppose that for some $m > 2$, it holds that any convex combination of m vectors lies in C. Let $\{x_1, \ldots, x_m, x_{m+1}\} \subseteq C$ and $\lambda_1, \ldots, \lambda_m, \lambda_{m+1} \geq 0$ such that $\sum_{i=1}^{m+1} \lambda_i = 1$. Our goal is to show that

$$z := \lambda_1 x_1 + \cdots + \lambda_m x_m + \lambda_{m+1} x_{m+1} \in C.$$

Observe there must exist at least on $\lambda \in [0, 1)$ or else if all $\lambda_i = 1$ the sum would be greater than 1. WLOG, assume that $\lambda_{m+1} \in [0, 1)$. Now

$$\begin{aligned} z = \sum_{i=1}^{m+1} \lambda_i x_i &= \left(\sum_{i=1}^{m} \lambda_i x_i \right) + \lambda_{m+1} x_{m+1} \\ &= (1 - \lambda_{m+1}) \left(\sum_{i=1}^{m} \frac{\lambda_i}{1 - \lambda_{m+1}} x_i \right) + \lambda_{m+1} x_{m+1} \\ &= (1 - \lambda_{m+1}) \left(\sum_{i=1}^{m} \lambda_i' x_i \right) + \lambda_{m+1} x_{m+1} \end{aligned}$$

Observe that $\lambda_i' := \dfrac{\lambda_i}{1 - \lambda_{m+1}} \geq 0$ and $\displaystyle\sum_{i=1}^{m} \lambda_i' = \frac{\lambda_1 + \cdots + \lambda_m}{1 - \lambda_{m+1}} = \frac{1 - \lambda_{m+1}}{1 - \lambda_{m+1}} = 1$ as $\displaystyle\sum_{i=1}^{m+1} \lambda_i = 1$. Then z is a convex combination of two vectors in C, so it also lies in C i.e.,

$$z = (1 - \lambda_{m+1}) \underbrace{\left(\sum_{i=1}^{m} \lambda_i' x_i \right)}_{\in C \text{ by IH}} + \lambda_{m+1} \underbrace{x_{m+1}}_{\in C} \in C.$$

It follows that C is convex as desired. ∎

Definition 1.7. Let $S \subseteq \mathbb{R}^n$. The intersection of all convex sets containing S is called the **convex hull** of S and is denoted by $\operatorname{conv}(S)$. It is the smallest convex set containing S.

Theorem 1.8. *Let $S \subseteq \mathbb{R}^n$. Then $conv(S)$ consists of all convex combinations of the elements of S, i.e.,*

$$conv(S) = \left\{ \sum_{i \in I} \lambda_i x_i \mid I \text{ is a finite index set and } \forall i \in I : x_i \in S, \lambda_i \geq 0, \sum_{i \in I} \lambda_i = 1 \right\}.$$

Proof. Define

$$D := \left\{ \sum_{i \in I} \lambda_i x_i \mid I \text{ is a finite index set and } \forall i \in I : x_i \in S, \lambda_i \geq 0, \sum_{i \in I} \lambda_i = 1 \right\}.$$

We first show $\operatorname{conv}(S) \subseteq D$. Clearly, $S \subseteq D$. Moreover, D is convex. Indeed, let $d_1, d_2 \in D$ and $\lambda \in [0,1]$. Then we can write

$$\begin{aligned} d_1 &= \sum_{i=1}^{k} \lambda_i x_i \quad \text{where} \quad \lambda_1, \ldots, \lambda_k \geq 0, \sum_{i=1}^{k} \lambda_i = 1, \{x_1, \ldots, x_k\} \subseteq S, \\ d_2 &= \sum_{j=1}^{r} \mu_j y_j \quad \text{where} \quad \mu_1, \ldots, \mu_r \geq 0, \sum_{j=1}^{r} \mu_j = 1, \{y_1, \ldots, y_r\} \subseteq S. \end{aligned}$$

Therefore,

$$\lambda d_1 + (1-\lambda) d_2 = [\lambda \lambda_1 x_1 + \cdots + \lambda \lambda_k x_k] + [(1-\lambda)\mu_1 y_1 + \cdots + (1-\lambda)\mu_r y_r].$$

Observe that $\lambda \lambda_1$ and $(1-\lambda)\mu_j$ are non-negative for all $i \in [k]$ and $j \in [r]$ and that

$$\begin{aligned} \lambda\lambda_1 + \cdots + \lambda\lambda_k + (1-\lambda)\mu_1 + \cdots + (1-\lambda)\mu_r &= \lambda \sum_{i=1}^{k} \lambda_i + (1-\lambda) \sum_{j=1}^{r} \mu_j \\ &= \lambda \cdot 1 + (1-\lambda) \cdot 1 = 1. \end{aligned}$$

Altogether, we conclude that D is a convex set $\supseteq S$ and thus $\operatorname{conv}(S) \subseteq D$.

We now show that $D \subseteq \operatorname{conv}(S)$. Observe that $S \subseteq \operatorname{conv}(S)$. Now combine with Theorem 1.6 to learn that the convex combinations of elements of S lie in $\operatorname{conv}(S)$. $\square$