1.2.2 Affine Sets and Affine Hull

Given $a, b \in \mathbb{R}^n$, the line connecting them is defined as

$$\mathcal{L}[a,b] := \{\lambda a + (1-\lambda)b |\ \lambda \in \mathbb{R}\}$$

Note that there is no restriction on λ.

Definition:(Affine Set) A subset S of $\mathbb{R}^n$ is *affine* if for any $a, b \in S$, we have $\mathcal{L}[a,b] \subseteq S$.

Definition:(Affine Combination)
Given $x_1, ..., x_m \in \mathbb{R}^n$, an element in the form $x = \sum_{i=1}^m \lambda_i x_i$, where $\sum_{i=1}^m \lambda_i = 1$ is called an affine combination of $x_1, ..., x_m$.

Proposition: A set S is affine if and only if it contains all affine combination of its elements.

Definition:(Affine Hull) The *affine hull* of a set $X \subseteq \mathbb{R}^n$ is

$$\text{aff}(X) := \bigcap \{S |\ S \text{ is affine and } X \subseteq S\}$$

Proposition: For any subset X of $\mathbb{R}^n$,

$$\text{aff}(X) = \Big\{ \sum_{i=1}^m \lambda_i x_i |\ \sum_{i=1}^m \lambda_i = 1,\ x_i \in X \Big\}$$

In fact, an *affine set* $S \subset \mathbb{R}^n$ is of the form $x + V$, where $x \in S$ and V is a vector space called the subspace parallel to S.

Lemma: Let S be nonempty. Then the following are equivalent:

1. S is affine
2. S is of the form $x + V$ for some subspace $V \subset \mathbb{R}^n$ and $x \in S$.

Also, V is unique and equals to $S - S$.

Proof. Suppose S is affine. We first assume $0 \in S$. Let $x \in S$ and $\gamma \in \mathbb{R}$. Since $0 \in S$, we have $\gamma x + (1-\gamma)0 = \gamma x \in S$. Now, suppose $x, y \in S$. Then $x + y = 2(\frac{1}{2}x + \frac{1}{2}y) \in S$. Hence, S is closed under addition and scalar multiplication. Therefore, $S = 0 + S$ is a linear subspace. If $0 \notin S$, then $0 \in S - x$ for any $x \in S$. So $S - x$ is a linear subspace. Therefore, $S = x + V$.
The other direction is simple, just use the fact that V is a linear subspace.
Now suppose $S = x_1 + V_1 = x_2 + V_2$, where $x_1, x_2 \in S$, V_1, V_2 are linear

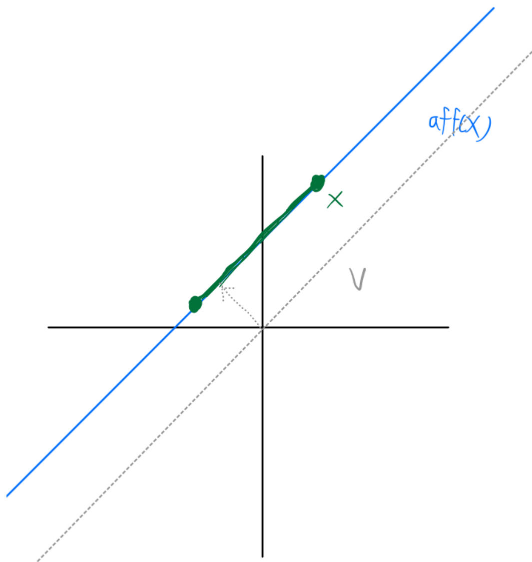

Figure 1: Affine hull and the parallel subspace

subspaces. Then $x_1 - x_2 + V_1 = V_2$. Since V_2 is a subspace, $x_1 - x_2 \in V_1$. So $V_2 = x_1 - x_2 + V_1 \subseteq V_1$. Similarly, $V_1 \subseteq V_2$. Therefore V is unique.
Since $S = x + V$, so $V = S - x \subseteq S - S$. Let $u, v \in S$ and $z = u - v$. Then $S - v = V$ by the uniqueness of V. So $z \in S - v = V$ and hence $S - S \subseteq V$. $\square$

Definition:(Dimension of affine and convex sets) The dimension of $\text{aff}(X)$ is defined to be the dimension of the subspace parallel to X. The dimension of a convex set C is defined to be the dimension of $\text{aff}(C)$.

Definition:(Affinely Independent) $x_0, ..., x_m \in \mathbb{R}^n$ are affinely independent if

$$\left[\sum \lambda_i x_i = 0, \ \sum \lambda_i = 0\right] \Longrightarrow [\lambda_i = 0 \text{ for all } i]$$

Proposition: $x_0, ..., x_m \in \mathbb{R}^n$ are affinely independent if and only if $x_1 - x_0, ..., x_m - x_0$ are linearly independent.

Proof. Suppose $x_0, ..., x_m$ are affinely independent. Suppose

$$\sum_{i=1}^{m} \lambda_i (x_i - x_0) = 0$$

Let $\lambda_0 := -\sum_{i=1}^{m} \lambda_i$, then we have

$$\lambda_0 x_0 + \sum_{i=1}^{m} \lambda_i x_i = 0$$

Since $\sum_{i=0}^{m} \lambda_i = 0$, $\lambda_i = 0$ for all i. Hence, $x_1 - x_0, ..., x_m - x_0$ are linearly independent.
The converse follows directly from the definition $\square$

Lemma: Let $S := \text{aff}(\{x_0, ..., x_m\})$, where $x_i \in \mathbb{R}^n$. Then $\text{span}\{x_1 - x_0, ..., x_m - x_0\}$ is the subspace parallel to S.

Proof. Let V be the subspace parallel to S. Then $S - x_0 = V$.
Hence $\text{span}\{x_1 - x_0, ..., x_m - x_0\} \subseteq V$.
Let $x \in V$, then $x + x_0 \in S$. So

$$x + x_0 = \sum_{i=0}^{m} \lambda_i x_i, \text{ where } \sum \lambda_i = 1$$

Therefore

$$x = \sum_{i=1}^{m} \lambda_i (x_i - x_0) \in \text{span}\{x_1 - x_0, x_m - x_0\}$$

$\square$

Proposition: $x_0, ..., x_m$ are affinely independent in $\mathbb{R}^n$ if and only if its affine hull is m-dimensional.

Proof. Suppose $x_0, ..., x_m$ are affinely independent. Then $x_1 - x_0, ..., x_m - x_0$ are linearly independent. Therefore, $V = \text{span}\{x_1 - x_0, ..., x_m - x_0\}$ is m-dimensional. Since V is the subspace parallel to $\text{aff}(\{x_0, ..., x_m\})$, $\text{aff}(\{x_0, ..., x_m\})$ is m-dimensional.
The converse is proven similarly. $\square$

Definition:(m-Simplex) Let $x_0, ..., x_m$ be affinely independent in $\mathbb{R}^n$. Then the set

$$\Delta_m := \text{conv}(\{x_0, ..., x_m\})$$

is called a m-simplex in $\mathbb{R}^n$ with vertices x_i.

Proposition: Consider a m-simplex Δ_m with vertices $x_0, ..., x_m$. For every $x \in \Delta_m$, there is a unique element $(\lambda_0, ..., \lambda_m) \in \mathbb{R}_+^{m+1}$ such that

$$x = \sum \lambda_i x_i, \; \sum \lambda_i = 1.$$

Proof. The existence follows directly from the definition. We only need to show the uniqueness.
Suppose $(\lambda_0, ..., \lambda_m),\ (\mu_0, ..., \mu_m) \in \mathbb{R}_+^{m+1}$ satisfy

$$x = \sum \lambda_i x_i = \sum \mu_i x_i,\ \sum \lambda_i = \sum \mu_i = 1$$

Then

$$\sum (\lambda_i - \mu_i) x_i = 0,\ \sum (\lambda_i - \mu_i) = 0$$

Since $x_0, ..., x_m$ are affinely independent, $\lambda_i - \mu_i = 0$ for all i. □

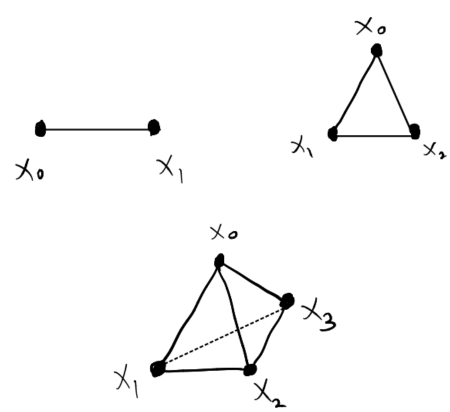

Figure 2: Examples of m-simplex

Definition: The cone generated by a set X is the set of all nonnegative combination of elements in X. A nonnegative (positive) combination of $x_1, x_2, ..., x_m$ is of the form

$$\sum_{i=1}^{m} \lambda_i x_i, \text{ where } \lambda_i \geq 0 \ (\lambda_i > 0).$$

Next, we prove a important theorem concerning convex hulls.

Theorem:(Caratheodory's Theorem) Let X be a nonempty subset of $\mathbb{R}^n$.

(a) Every nonzero vector of cone(X) can be represented as a positive combination of linearly independent vectors from X.

(b) Every vector from conv(X) can be represented as a convex combination of at most $n+1$ vectors from X.

Proof. (a) Let $x \in \text{cone}(X)$ and $x \neq 0$. Suppose m is the smallest integer such that x is of the form $\sum_{i=1}^{m} \lambda_i x_i$, where $\lambda_i > 0$ and $x_i \in X$. Suppose that x_i are not linearly independent. Therefore, there exist μ_i with at least one μ_i positive, such that $\sum_{i=1}^{m} \mu_i x_i = 0$. Consider $\overline{\gamma}$, the largest γ such that $\lambda_i - \gamma\mu_i \geq 0$ for all i. Then $\sum_{i=1}^{m} (\lambda_i - \overline{\gamma}\mu)x_i$ is a representation of x as a positive combination of less than m vectors, contradiction. Hence, x_i are linearly independent.
(b) Consider $Y = \{(x, 1) : x \in X\}$. Let $x \in \text{conv}(X)$. Then $x = \sum_{i=1}^{m} \lambda_i x_i$, where $\sum_{i=1}^{m} \lambda_i = 1$, so $(x, 1) \in \text{cone}(Y)$.
By (a), $(x, 1) = \sum_{i=1}^{l} \lambda'_i (x_i, 1)$, where $\lambda_i > 0$. Also, $(x_1, 1), ..., (x_l, 1)$ are linearly independent vectors in $\mathbb{R}^{n+1}$(at most $n + 1$). Hence, $x = \sum_{i=1}^{l} \lambda'_i x_i$, $\sum_{i=1}^{m} \lambda'_i = 1$ □

Proposition: Let $X \subseteq \mathbb{R}^n$ be a compact set. Then $\text{conv}(X)$ is compact.

Proof. Let $\{x^k\}$ be a sequence in $\text{conv}(X)$. By Caratheodory's Theorem,

$$x^k = \sum_{i=1}^{n+1} \lambda_i^k x_i^k$$

where $\lambda_i^k \geq 0$, $x_i^k \in X$ and $\sum_{i=1}^{n+1} \lambda_i^k = 1$.
Note that the sequence $\{(\lambda_1^k, ..., \lambda_{n+1}^k, x_1^k, ..., x_{n+1}^k)\}$ is bounded. Then it has a limit point $(\lambda_1, ..., \lambda_{n+1}, x_1, ..., x_{n+1})$, where $\sum_{i=1}^{n+1} \lambda_i = 1$ and $x_i \in X$.
Hence $\sum_{i=1}^{n+1} \lambda_i x_i \in \text{conv}(X)$ is a limit point of the sequence x^k.
Therefore, $\text{conv}(X)$ is compact. □

1.3 Convex Functions

In this course, we will consider extended-real-valued functions, which take values in $\overline{\mathbb{R}} := (-\infty, \infty]$, with the convention that $a + \infty = \infty \ \forall a \in \mathbb{R}$, $\infty + \infty = \infty$, and $t \cdot \infty = \infty \ \forall t > 0$.

1.3.1 Convex Functions

Definition:(Convex Functions) Let C be a convex subset of $\mathbb{R}^n$. A function $f : C \to \overline{\mathbb{R}}$ is called *convex* on C if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)y, \forall x, y \in C, \forall \lambda \in [0, 1].$$

A function is called *stricly convex* if the inequality above is strict for all $x, y \in C$ with $x \neq y$, and all $\lambda \in (0, 1)$. A function is called *concave* if $(-f)$ is convex.

Definition:(Level Sets) For a function $f : C \to \mathbb{R}$, we define the *level sets* of f to be $\{x \mid f(x) \leq \lambda\}$.

If a function is convex, then all its level sets are also convex (Exercise).

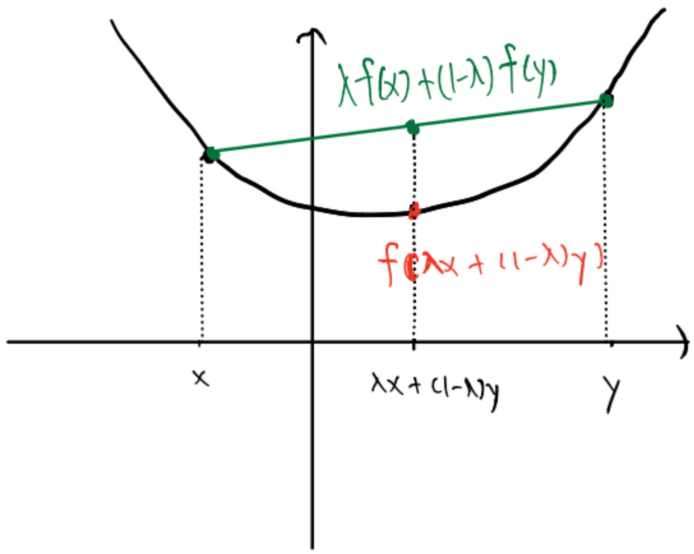

Figure 3: Convex Function

However, the convexity of all level sets of a function does not necessarily imply the convexity of the function itself.

Examples of Convex Functions
The following functions are convex:

(a) $f(x) := \langle a, x\rangle + b$ for $x \in \mathbb{R}^n$, where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

(b) $g(x) := ||x||$ for $x \in \mathbb{R}^n$.

(c) $h(x) := x^2$ for $x \in \mathbb{R}$.

(d) $F(x) := \frac{1}{2}x^T Ax$ for $x \in \mathbb{R}^n$, where A is a $n \times n$ symmetric positive semidefinite matrix. (i.e. $x^T Ax \geq 0$ for all $x \in \mathbb{R}^n$)

Definition:(Epigraph and Effective Domain)
The *epigraph* of a function $f : X \to [-\infty, \infty]$, where $X \subset \mathbb{R}^n$, is given by

$$\text{epi} f = \{(x, w) |\ x \in X,\ w \in \mathbb{R},\ f(x) \leqslant w\}.$$

The *effective domain* of f is given by

$$\text{dom} f = \{x |\ f(x) < \infty\}.$$

Note that dom$f$ is just the projection of epi$f$ on $\mathbb{R}^n$.

Definition:(Proper Function)
A function f is *proper* if $f(x) < \infty$ for at least one $x \in X$. f is *improper* if it is not proper. By considering epi$f$, f is proper means that epi$f$ is not empty and does not contain any vertical line.

Theorem:(Jensen inequality)
A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if and only if for any $\lambda_i \geq 0$ with $\sum \lambda_i = 1$ and for any elements $x_i \in \mathbb{R}^n$, it holds that

$$f\big(\sum \lambda_i x_i\big) \leq \sum \lambda_i f(x_i)$$

Proof. It suffices to prove that any convex function satisfies the Jensen inequality. We will prove this by induction.
The case $m = 1, 2$ are simple. So suppose the inequality holds for all $k \leq m$.
Suppose $\lambda_i \geq 0$ satisfies $\sum_{i=1}^{m+1} \lambda_i = 1$. Then $\sum_{i=1}^{m} \lambda_i = 1 - \lambda_{m+1}$.
If $\lambda_{m+1} = 1$, then $\lambda_i = 0$ for all i. Then the inequality holds.
So suppose $\lambda_{m+1} < 1$. Then

$$\sum_{i=1}^{m} \frac{\lambda_i}{1 - \lambda_{m+1}} = 1$$

and

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

□

The following gives a geometric characterization of convexity.

Proposition: A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if and only if epi$f \subset \mathbb{R}^{n+1}$ is convex.

Proof. Assume f is convex. Let $(x_1, t_1), (x_2, t_2) \in$ epi$f$ and $\lambda \in [0, 1]$. Then

$$f(\lambda x_1 + (1 - \lambda) x_2) \leq \lambda f(x_1) + (1 - \lambda) f(x_2) \leq \lambda t_1 + (1 - \lambda) t_2$$

Hence $(\lambda(x_1, t_1) + (1-\lambda)(x_2, t_2) \in \text{epi}f$.
Conversely, suppose epi$f$ is convex. Let $x_1, x_2 \in \text{dom}f$ and $\lambda \in [0,1]$.
Since epi$f$ is convex, $\lambda(x_1, f(x_1)) + (1-\lambda)(x_2, f(x_2)) \in \text{epi}f$. Then

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

Therefore, f is convex. □

Definition:(Closed function) If the epigragh of a function $f : X \to \overline{\mathbb{R}}$ is closed, we say that f is a *closed* function.

For example, the indicator funtion δ_X is convex if and only if X is convex, is closed if and only if X is closed, where

$$\delta_X(x) := \begin{cases} 0 & x \in X \\ \infty & \text{otherwise} \end{cases}$$

In fact, closedness is related to the concept of lower semicontinuity.
Recall that a function f is called *lower semicontinuous* at $x \in X$ if

$$f(x) \leq \liminf_{k\to\infty} f(x_k)$$

for every sequence $\{x_k\} \subset X$ with $x \to x_k$. f is *lower semicontinuous* if it is lower semicontinuous at each $x \in X$. f is *upper semicontinuous* if $-f$ is lower semicontinuous.

Proposition: Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function, then the following are equivalent:

(i) The level set $V_\gamma = \{x | f(x) \leq \gamma\}$ is closed for every γ.

(ii) f is lower semicontinuous.

(iii) epi$f$ is closed.

Proof. If $f(x) = \infty$ for all x, then the result holds. So assume $f(x) < \infty$ for some $x \in \mathbb{R}^n$. Therefore, epi$f$ is nonempty and there exists level sets of f that are nonempty.
(i) $\implies$ (ii). Assume V_γ is closed for every γ. Suppose f is not lower semicontinuous, that is

$$f(x) > \liminf_{k\to\infty} f(x_k)$$

for some x and sequence $\{x_k\}$ converging to x. Let γ satisfies

$$f(x) > \gamma > \liminf_{k\to\infty} f(x_k).$$

Hence, there exists a subsequence $\{x_{k_i}\}$ such that $f(x_{k_i}) \leq \gamma$ for all i. So, $\{x_{k_i}\} \subset V_\gamma$. But V_γ is closed, x also belongs to V_γ. Therefore, $f(x) \leq \gamma$, contradiction.

(ii) $\implies$ (iii). Assume f is lower semicontinuous. Let (x, w) be the limit of $\{(x_k, w_k)\} \subset \text{epi}(f)$. We have $f(x_k) \le w_k$ for all k. Since f is lower semicontinuous, taking limit we have,

$$f(x) \le \liminf_{k \to \infty} f(x_k) \le w.$$

Hence $(x, w) \in \text{epi}f$ and so $\text{epi}f$ is closed.
(iii) $\implies$ (i). Assume $\text{epi}f$ is closed. Let $\{x_k\}$ be a sequence in V_γ converging to x for some γ. We have $f(x_k) \le \gamma$, so $(x_k, \gamma) \in \text{epi}f$ for each k. Since $\text{epi}f$ is closed and $(x_k, \gamma) \to (x, \gamma)$, we have $(x, \gamma) \in \text{epi}f$, that is $f(x) \le \gamma$. Hence $x \in V_\gamma$ and V_γ is closed. □