1.2 Convex and Affine Hulls

1.2.1 Convex Hull

Definition:(Convex Hull)
Let X be a subset of $\mathbb{R}^n$. The convex hull of X is defined by

$$\text{conv}(X) := \bigcap \{C|\ C \text{ is convex and } X \subseteq C\}$$

In other words, conv(X) is the smallest convex set containing X.
The next proposition provides a good representation for elements in the convex hull.

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

$$\text{conv}(X) = \Big\{ \sum_{i=1}^{m} \lambda_i x_i |\ \sum_{i=1}^{m} \lambda_i = 1,\ \lambda_i \geq 0,\ x_i \in X \Big\}$$

Proof. Let $C = \big\{ \sum_{i=1}^{m} \lambda_i x_i |\ \sum_{i=1}^{m} \lambda_i = 1,\ \lambda_i \geq 0,\ x_i \in X \big\}$. Clearly, $X \subseteq C$. Next, we check that C is convex.
Let $a = \sum_{i=1}^{p} \alpha_i a_i, b = \sum_{j=1}^{q} \beta_j b_j$ be elements of C, where $a_i, b_i \in C$ with $\alpha_i, \beta_j \geq 0$ and $\sum \alpha_i = \sum \beta_j = 1$. Suppose $\lambda \in [0,1]$, then

$$\lambda a + (1-\lambda) b = \sum_{i=1}^{p} \lambda \alpha_i a_i + \sum_{j=1}^{q} (1-\lambda) \beta_j b_j.$$

Since

$$\sum_{i=1}^{p} \lambda \alpha_i + \sum_{j=1}^{q} (1-\lambda) \beta_j = \lambda \sum_{i=1}^{p} \alpha_i + (1-\lambda) \sum_{j=1}^{q} \beta_j = 1$$

we have $\lambda a + (1-\lambda) b \in C$. Hence, C is convex. Also, conv(X) $\subseteq C$ by the definition of conv(X).
Suppose $a = \sum \lambda_i a_i \in C$. Then since each $a_i \in X \subseteq$ conv(X) and conv(X) is convex, we have $a \in$ conv(X). Therefore, conv(X) $=C$. $\square$

Let $a, b \in \mathbb{R}^n$, define the interval

$$[a, b) := \{\lambda a + (1-\lambda) b|\ \lambda \in (0, 1]\}$$

The intervals $(a, b], (a, b)$ are defined similarly.

Lemma: For a convex set $C \in \mathbb{R}^n$ with nonempty interior, take $a \in C^{\text{o}}$ and $b \in \overline{C}$. Then $[a, b) \subset C^{\text{o}}$.

Proof. Since $b \in \overline{C}$, for any $\epsilon > 0$, we have $b \in C + \epsilon \mathbf{B}$, where $\mathbf{B}$ denotes the closed unit ball centered at 0.

Take $\lambda \in (0,1]$ and let $x_\lambda := \lambda a + (1-\lambda)b$. Let ϵ be such that $a + \epsilon\frac{2-\lambda}{\lambda}\mathbf{B} \subset C$.

$$\begin{aligned} x_\lambda + \epsilon\mathbf{B} &= \lambda a + (1-\lambda)b + \epsilon\mathbf{B} \\ &\subset \lambda a + (1-\lambda)[C + \epsilon\mathbf{B}] + \epsilon\mathbf{B} \\ &= \lambda a + (1-\lambda)C + (2-\lambda)\epsilon\mathbf{B} \\ &\subset \lambda[a + \epsilon\frac{2-\lambda}{\lambda}\mathbf{B}] + (1-\lambda)C \\ &\subset \lambda C + (1-\lambda)C \subset C \end{aligned}$$

Hence $x_\lambda \in C^o$ and $[a,b) \subset C^o$.

□

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 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 $S - S \subseteq V$. □

Definition:(Dimension of affine and convex sets) The dimension of 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 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 □

Lemma: Let $S :=$ aff$(\{x_0, ..., x_m\})$, where $x_i \in \mathbb{R}^n$. Then 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 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

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

□

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 aff$(\{x_0, ..., x_m\})$, aff$(\{x_0, ..., x_m\})$ is m-dimensional.
The converse is proven similarly. □

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. □

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 $\text{cone}(X)$ can be represented as a positive combination of linearly independent vectors from X.

(b) Every vector from $\text{conv}(X)$ 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. □