I. CONVEX OPTIMIZATION

In convex optimization, our goal is to minimize an objective a convex function subject to convex inequality and affine equality constraints. The problem can be mathematically written as:

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq b_i \quad i = 1, \ldots, k \\
& \quad g_j(x) = 0 \quad j = 1, \ldots, l
\end{align*}
\]

where \( f_0(x) \) and \( \{f_i(x)\}_{i=1}^k \) are convex functions, and \( \{g_j(x)\}_{j=1}^l \) are affine.

Throughout the following lectures, the following notations and definitions will be used:

A. Notations

1) \( \mathbb{R} \) - the set of all real numbers
2) \( \mathbb{R}_+ \) - the set of all non-negative real numbers
3) \( \mathbb{R}_{++} \) - the set of all positive real numbers
4) \( \mathbb{R}^n \) - the set of all real \( n \)-dimensional vectors
5) \( \mathbb{R}^{m \times n} \) - the set of real \( m \times n \) matrices
6) \( (a, b, c) \) - a column vector whose elements (by order) are \( a, b \) and \( c \).
7) \( (\cdot)^T \) - the transpose operator.
8) \( \mathbf{1} \) - a column vector composed of one’s: \( (1, 1, \ldots, 1) \)
9) \( x \) - a vector
10) \( x_i \) - the \( i^{th} \) element of \( x \).
11) \( S^k \) - the set of all (real) symmetric \( k \times k \) matrices: \( S^k = \{ A \in \mathbb{R}^{k \times k} : A^T = A \} \)
12) \( S^+_k \) - the set of all (real) symmetric positive semi-definite (PSD) \( k \times k \) matrices:
    \[ S^+_k = \{ A \in S^{k \times k} : S \succeq 0 \} \]
13) \( S^{++}_k \) - the set of all (real) symmetric positive semi-definite (PSD) \( k \times k \) matrices:
    \[ S^{++}_k = \{ A \in S^{k \times k} : S \succ 0 \} \]
14) \( x \succeq y \) (vectors) - element-wise inequality: \( \forall i \quad x_i \geq y \)
15) \( A \succeq B \) (matrices) - \( (A - B) \in S^+_k \)
16) \( \text{dom}(f) \) - the domain on which \( f \) is defined. For example: \( \log : \mathbb{R} \to \mathbb{R} \), and
    \[ \text{dom}(f) = \mathbb{R}_{++} \]
17) \( \bar{\theta} = 1 - \theta \)

Note the difference between the interpretations of inequalities in \( \mathbb{R}^n \) when \( n = 1 \) and \( n > 1 \). As \( \mathbb{R} \) is well-ordered, the elements \( a, b \in \mathbb{R} \) satisfy \( a < b, a = b \) or \( a > b \). However, for \( \mathbb{R}^2 \) the vectors \((0, 1)\) and \((1, 0)\) do not satisfy any of the relations \( <, =, > \).
B. Definitions

Let \( x_1, x_2 \in \mathbb{R}^n \) be two arbitrary vectors. Then:

1) The **line** that goes through \( x_1 \) and \( x_2 \) is the set of all points \( \theta x_1 + \bar{\theta} x_2 \), where \( \theta \in \mathbb{R} \).

   ![A line](Fig. 1. A line)

2) The **line segment** that goes through \( x_1 \) and \( x_2 \) is the set of all points \( \theta x_1 + \bar{\theta} x_2 \), where \( \theta \in [0, 1] \).

   ![A line segment](Fig. 2. A line segment)

3) A set \( C \subseteq \mathbb{R}^n \) is an **affine set** if any line that goes through any two points in \( C \) is contained in \( C \):

   \[
   \forall \theta \in \mathbb{R} : \quad x_1, x_2 \in C \Rightarrow \theta x_1 + \bar{\theta} x_2 \in C.
   \]

4) A set \( C \subseteq \mathbb{R}^n \) is a **convex set** if the line segment that goes through any two points in \( C \) is contained in \( C \):

   \[
   \forall \theta \in [0, 1] : \quad x_1, x_2 \in C \Rightarrow \theta x_1 + \bar{\theta} x_2 \in C.
   \]

   For example, see Figure 3.

   ![To the left: a convex set. To the right: a non-convex set.](Fig. 3)

5) A **convex combination** of \( (x_1, x_2, \ldots, x_k) \) is any linear combination \( \sum_{i=1}^{k} \theta_i x_i \), where \( \theta_i \geq 0 \) and \( \sum_{i=1}^{k} \theta_i = 1 \).

6) A **convex hull**, denoted \( \text{Conv}(C) \), is the set of all convex combinations of points in \( C \). A convex hull is the smallest convex set that contains \( C \). For example, in Figure 4, the convex hull of the non-convex set is obtained by adding the area inside the dashed line.

   ![To the left: a convex set. To the right: a non-convex set.](Fig. 4)
7) A set $C$ is called a **cone** if $\forall x \in C$ and $\theta \geq 0$ we have $\theta x \in C$.

8) A cone $C$ is called a **convex cone** if it is convex: 
   $\forall \theta_1, \theta_2 \geq 0 : \ x_1, x_2 \in C \Rightarrow \theta_1 x_1 + \theta_2 x_2 \in C$.

9) A **conic combination** of $(x_1, x_2, \ldots, x_x)$ is any linear combination $\sum_{i=1}^{k} \theta_i x_i$, where $\theta_i \geq 0$.

10) A **conic hull** is the set of all conic combinations of points in $C$.

11) A cone $C$ is called **pointed** if $x \in C$ and $-x \in C$ implies that $x = 0$. This means that the cone $C$ contains no line.

12) A cone $C \subset \mathbb{R}^n$ is called a **proper cone** if the following requirements hold:
   a) $C$ is a convex cone.
   b) $C$ is a closed set.
   c) $C$ has a non-empty interior.
   d) $C$ is pointed.

13) A **ray** is a set of points $\{ x : x = x_0 + \theta y | \theta \geq 0 \}$.

14) A **Hyperplane** is a set $\{ x : a^T x = b \}$, where $a, x \in \mathbb{R}^n$, $a \neq 0$ and $b \in \mathbb{R}$.
   Another possible notation is $\{ x : a^T(x - x_0) = b \}$, where $x_0$ is any point on the hyperplane (so that $a^T x_0 = b$).
a is called the **normal** to the hyperplane.

15) A **Half-space** is the set of points above (or below) a hyperplane:
\[ \{ x : a^T (x - x_0) > b \} \]

16) A **Polyhedra** is the set of points that satisfy a set of linear equalities & inequalities. For instance:

\[
P = \{ x \in \mathbb{R}^n \}
\]

\[
s.t. \begin{cases} 
a_i^T x \leq b_i & i = 1, 2, \ldots, K \\
c_j^T x = d_j & j = 1, 2, \ldots, J 
\end{cases}
\]

A compact notation is
\[
P = \{ x \in \mathbb{R}^n : Ax = b, Cx = d \}
\]

where \( A = (a_1^T, \ldots, a_K^T) \) and \( C = (c_1^T, \ldots, c_J^T) \).

A Polyhedra is the outcome of an intersection of half-spaces and hyperplanes.

**C. Examples**

1) The empty set \( \emptyset \), a single point in \( x_0 \in \mathbb{R}^n \) and the entire space are affine.
2) Any line is affine.
3) Any line that passes through the origin is a cone.
4) A line segment is convex, but not affine.
5) A ray is convex but not affine. It is a cone iff \( x_0 = 0 \).
6) A subspace is convex, affine & a cone.

We start with the following lemma regarding convex sets:

**Lemma 1** Convexity is preserved under intersections: Let \( S_1, S_2 \) be convex sets. Then \( S_1 \cap S_2 \) is convex.

**Exercise:** Prove Lemma 1.

Later we will see that if \( f_1(x) \) and \( f_2(x) \) are convex, then \( f(x) = \max(f_1(x), f_2(x)) \) is convex, and we will see that this is equivalent to Lemma 1.

## II. Generalized Inequalities

### A. Definition

Let \( K \subset \mathbb{R}^n \) be a proper cone. We define the generalized inequality with respect to \( K \) as following:

\[ x \preceq_K y \iff x - y \in K \quad (4) \]

### B. Properties of the Generalized Inequality

1) The G.I. is preserved under addition: \( x_1 \preceq_K y_1, \ x_2 \preceq_K y_2 \Rightarrow x_1 + x_2 \preceq_K y_1 + y_2 \)
2) The G.I. is transitive: \( a \preceq_K b, \ b \preceq_K c \Rightarrow a \preceq_K c \)
3) The G.I. is preserved under non-negative scaling: \( x \preceq_K y, \ a \geq 0 \Rightarrow ax \preceq_K ay \)
4) The G.I. is reflexive: \( \forall x : \ x \preceq_K x \)
5) The G.I. is preserved under limits: if \( \forall i \ x_i \preceq_K y_i \), and \( x = \lim_{i \to \infty} x_i \) and \( y = \lim_{i \to \infty} y_i \) exist, then \( x \preceq_K y \)

**Exercise:** Prove properties 1-5.

**Theorem 1** Separation Theorem

Let \( C, D \) be two convex sets in \( \mathbb{R}^n \), such that \( C \cap D = \emptyset \). Then, there exists a hyperplane \( a^T (x - x_0) = 0 \) such that:

- \( \forall x \in C : \ a^T (x - x_0) \leq 0 \), and
- \( \forall x \in D : \ a^T (x - x_0) \geq 0 \).

**Proof:** Assume that \( C, D \) are closed sets. Define the following distance measure:

\[ \text{dist} (C, D) \triangleq \min_{u \in C, v \in D} \| u - v \|^2 \quad (5) \]

Let \( c \in C, d \in D \) be the elements that achieve the minimum:

\[ \| c - d \|^2 = \text{dist} (C, D) = \min_{u \in C, v \in D} \| u - v \|^2 \quad (6) \]

and let \( a \) be the line connecting \( c, d \) \((a = d - c)\), and \( x_0 = \frac{c + d}{2} \). Define the following functional:

\[ f(x) = a^T (x - x_0) = a^T \left( x - \frac{c + d}{2} \right) = (d - c)^T \left( x - \frac{c + d}{2} \right) \quad (7) \]
\[ (d - c)^T (x - d + \frac{d - c}{2}) = (d - c)^T (x - d) + \frac{1}{2} \|d - c\|^2 \]

We show that for every \( u \in D \) we have \( f(u) \geq 0 \):
Assume that there exists a vector \( u \in D \) s.t. \( f(u) < 0 \). Then we may write
\[
0 > f(u) = (d - c)^T (u - d) + \frac{1}{2} \|d - c\|^2 > (d - c)^T (u - d) \tag{8}
\]
Define the following function:
\[
g(t) = \|d + t(u - d) - c\|^2 = \|(1 - t) d + tu - c\|^2 \tag{9}
\]
\[
= ((1 - t) d + t(u - d) - c)^T ((1 - t) d + tu - c)
\]
\[
= (d + t(u - d) - c)^T (d + tu - c)
\]
\[
\frac{d}{dt} g(t) = \frac{d}{dt} \left[ (d + t(u - d) - c)^T \right] (d + t(u - d) - c) \tag{10}
\]
\[
= (u - d)^T (d + t(u - d) - c) + (d + t(u - d) - c)^T (u - d)
\]
and inspect
\[
\frac{d}{dt} g(t)|_{t=0} = 2 (d - c)^T (u - d) \tag{11}
\]
By (8) we have that
\[
\frac{d}{dt} g(t)|_{t=0} < 0 \tag{12}
\]
so \( g(t) \) is decreasing at \( t = 0 \). If so, there exists a (small) \( t_0 \) s.t. \( 0 < t_0 < 1 \) for which \( g(t_0) < g(0) \), so
\[
\|(1 - t_0) d + t_0 u - c\|^2 < \|d - c\|^2 \tag{13}
\]
But since \( D \) is convex, and by assumption we have \( u, d \in D \), we also have \( u' = (1 - t_0) d + t_0 u \in D \). So, we may write
\[
\|u' - c\|^2 < \|d - c\|^2 = \text{dist} (C, D) \tag{14}
\]
which contradicts the assumption that \( c, d \) satisfy the minimum in (6). Therefore we conclude that \( f(u) \geq 0 \) for all \( u \in D \). The proof of \( f(v) \leq 0 \) for all \( v \in C \) follows from similar steps.