Convex Optimization - AI2101
Assignment - I

Bhuvan Chandra K
AI23BTECH11013

January 26, 2025

Question 2.1

Let $C \subseteq \mathbb{R}^n$ be a convex set, with $x_1, \dots, x_k \in C$, and let $\theta_1, \dots, \theta_k \in \mathbb{R}$ satisfy $\theta_i \geq 0$,

$$\theta_1 + \cdots + \theta_k = 1.$$

Show that

$$\theta_1 x_1 + \cdots + \theta_k x_k \in C.$$

(The definition of convexity is that this holds for $k = 2$; you must show it for arbitrary k.)
Hint. Use induction on k.

Solution:

Given, $C \subseteq \mathbb{R}^n$ be a convex set, with $x_1, x_2, \dots, x_k \in C$, and $\theta_1, \dots, \theta_k \in \mathbb{R}$ which satisfy $\theta_i \geq 0$ and $\theta_1 + \cdots + \theta_k = 1$.

To prove: $\theta_1 x_1 + \cdots + \theta_k x_k \in C$.

We will prove this result using induction.

1. Base case ($k = 1$):
   When $k = 1$, it is a single point $x_1 \in C$ and $\theta_1 = 1$, so $\theta_1 x_1 = x_1 \in C$.

   Now let's assume that the given statement holds for k points $x_1, \dots, x_k \in C$ and $\theta_1, \dots, \theta_k \geq 0$, also $\theta_1 + \cdots + \theta_k = 1$, which results in:

   $$\theta_1 x_1 + \cdots + \theta_k x_k \in C.$$

2. Induction step:
   We need to prove that this statement holds for the case with $k + 1$ points.

   Let's consider $k + 1$ points $x_1, \dots, x_k, x_{k+1} \in C$, and $\theta_1, \dots, \theta_k, \theta_{k+1} \geq 0$, and $\theta_1 + \cdots + \theta_k + \theta_{k+1} = 1$.

   Let $r = \theta_1 + \cdots + \theta_{k+1}$, where $r = 1$.

   $t = \theta_1 + \cdots + \theta_k$. We know $t + \theta_{k+1} = r = 1$.

   For the k points $x_1, \dots, x_k$, we define coefficients in the form

   $$y_i = \frac{\theta_i}{t}, \quad \forall i = 1, \dots, k.$$

   $$\sum_{i=1}^{k} y_i = 1 \quad \text{(since } \sum_{i=1}^{k} \theta_i = t\text{)}$$
   $$y_i \geq 0, \forall i, \quad \text{as } \theta_i \geq 0 \text{ and } t \geq 0$$

   $$y_i = \frac{\theta_i}{\sum_{i=1}^{k} \theta_i}.$$

Since the statement holds in the case of first k points $x_1, \dots, x_k$, we can write:

$$y_1 x_1 + \dots + y_k x_k \in C,$$

since $\sum y_i = 1$, $y_i \geq 0$.

Now:

$$\theta_1 x_1 + \dots + \theta_k x_k + \theta_{k+1} x_{k+1} = t(y_1 x_1 + \dots + y_k x_k) + \theta_{k+1} x_{k+1}.$$

As we know $t + \theta_{k+1} = 1$, So we can write,

$$\theta_1 x_1 + \dots + \theta_k x_k + \theta_{k+1} x_{k+1} = (1 - \theta_{k+1})(\sum_{i=1}^{k} y_i x_i) + (\theta_{k+1} x_{k+1})$$

Since,

$$\sum_{i=1}^{k} y_i x_i \in C, x_{k+1} \in C,$$

From definition of convexity (which is in this case(for two points) holds, this means the point is in C. That is,

$$\sum_{i=1}^{k+1} \theta_i x_i \in C$$

$\therefore$ By Induction, the given statement is true for all $k \geq 1$.

Question 2.4

Show that the convex hull of a set S is the intersection of all convex sets that contain S. (The same method can be used to show that the conic, or affine, or linear hull of a set S is the intersection of all conic sets, or affine sets, or subspaces that contain S.)

Solution:

Let H be the convex hull of S and let I be the intersection of all convex sets that contain S.

Note: Convex Hull is the smallest convex set containing S.

Let us show that $H = I$ by proving $H \subseteq I$ and $I \subseteq H$.

(i) $H \subseteq I$:

Let $x \in H$.

Since H is a convex set and contains S,
x is a convex combination:

x = $\sum \theta_i x_i$

where $\theta_i \geq 0 \, \forall i$, $\sum \theta_i = 1$, and $x_i \in S$ for $i = 1, 2, \dots, n$.

That is, since H contains S and is convex, the above linear combination x will be in H.

Now, I is the intersection of all convex sets containing S, which means I is convex and contains the points that are in S.

i.e., all of $x_i \in I$.

I is convex, so convex combination of points in I will be in I.

We can say $x \in I$ for all such convex sets.

$$x \in H \implies x \in I$$

Thus, $H \subseteq I$.

(ii) $I \subseteq H$:

I is the intersection of all convex sets containing S, which means

$$I = \bigcap \{C \mid C \text{ is convex and } S \subseteq C\}.$$

Suppose $y \in I$, y is in the intersection of all convex sets containing S.
This means that y must be in every one of those convex sets containing S to be in I.
H is also a convex set containing S (From the definition of Convex Hull), so $y \in H$.
Thus,

$$y \in I \implies y \in H.$$

So, $I \subseteq H$ Since $H \subseteq I$ and $I \subseteq H$, We can conclude that H = I.

$\therefore$ The convex hull of a set S is the intersection of all convex sets that contain S.

Question 2.5

What is the distance between two parallel hyperplanes $\{x \in \mathbb{R}^n \mid a^T x = b_1\}$ and $\{x \in \mathbb{R}^n \mid a^T x = b_2\}$?

Solution:

Given two parallel hyperplanes
Equations:

$$a^T x = b_1$$

$$a^T x = b_2$$

We can see that a is the normal vector to both hyperplanes.

b_1 and b_2 are constants that position the hyperplane in the space. As the shortest distance line between these planes must be perpendicular to both, i.e., the points must lie along the normal vector direction.

If not along the normal vector, then it wouldn't be the shortest distance. As the shortest distance line wouldn't be perpendicular to the respective plane. Let these points be x_1, lying on the plane $a^T x = b_1$, and x_2, lying on the plane $a^T x = b_2$.

We know,

$$x_1 = k_1 a, \quad x_2 = k_2 a$$

Since x_1 and x_2 are along normal vector , they can be written as a scalar multiple of the normal vector.

$$a^T(k_1 a) = b_1, \quad a^T(k_2 a) = b_2$$

$$a^T a = \|a\|_2^2$$

$$k_1 = \frac{b_1}{\|a\|_2^2}, \quad k_2 = \frac{b_2}{\|a\|_2^2}$$

Shortest distance = Distance between x_1 and x_2

$$= \|x_1 - x_2\|_2$$

$$\begin{aligned}
\|x_1 - x_2\|_2 &= \left\| \frac{b_1}{\|a\|_2^2} a - \frac{b_2}{\|a\|_2^2} a \right\| \\
&= \left\| \frac{b_1 - b_2}{\|a\|_2^2} a \right\| \\
&= \left| \frac{b_1 - b_2}{\|a\|_2^2} \right| \|a\| = \frac{|b_1 - b_2|}{\|a\|_2}
\end{aligned}$$

$$\therefore \text{ Distance between the hyperplanes } = \frac{|b_1 - b_2|}{\|a\|_2}.$$

Question 2.7

Voronoi description of halfspace. Let a and b be distinct points in $\mathbb{R}^n$. Show that the set of all points that are closer (in Euclidean norm) to a than b, i.e.,

$$\{x \mid \|x - a\|_2 \leq \|x - b\|_2\},$$

is a halfspace. Describe it explicitly as an inequality of the form $c^T x \leq d$. Draw a picture.

Solution:

Definition: A halfspace in $\mathbb{R}^n$ is a subset of the form

$$\{x \in \mathbb{R}^n \mid c^T x \leq d\}$$

where $c \in \mathbb{R}^n$ is a nonzero vector and $d \in \mathbb{R}$ is a scalar.

Let $a, b \in \mathbb{R}^n$ be distinct points.

Define the set S:

$$S = \{x \in \mathbb{R}^n \mid \|x - a\|_2 \leq \|x - b\|_2\}.$$

Squaring these does not change the inequality:

$$\|x - a\|_2^2 \leq \|x - b\|_2^2.$$

Expanding both sides:

$$(x - a)^T (x - a) \leq (x - b)^T (x - b).$$

Simplify:

$$x^T x - 2a^T x + a^T a \leq x^T x - 2b^T x + b^T b.$$

Cancel $x^T x$:

$$-2a^T x + a^T a \leq -2b^T x + b^T b.$$

Rearranging:

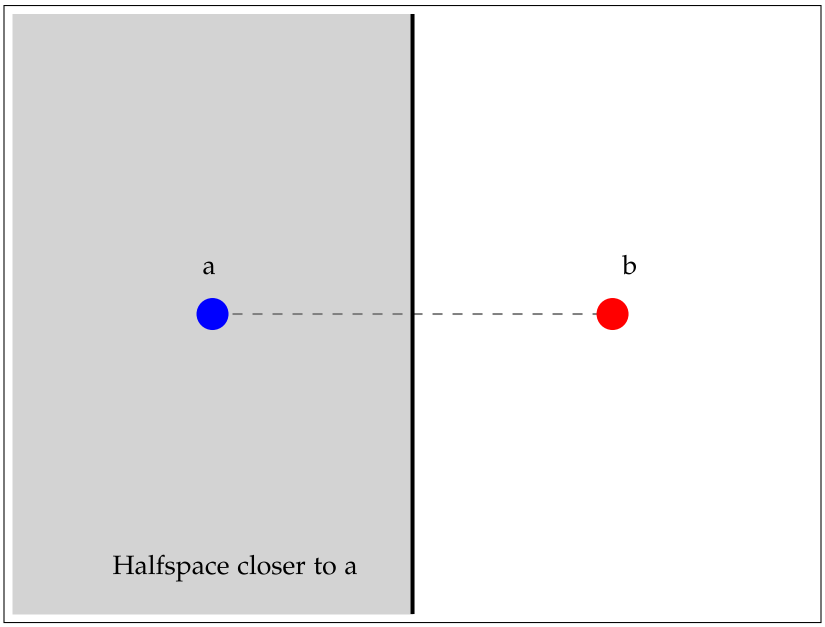
$$2(b-a)^T x \le b^T b - a^T a.$$

In halfspace representation:

$$c = 2(b-a), \quad d = b^T b - a^T a,$$
$$S = \{x \in \mathbb{R}^n \mid c^T x \le d\}.$$

$\therefore$ Given, set S is a halfspace.

Following picture depicts the halfspace:



Question 2.8

Which of the following sets S are polyhedra? If possible, express S in the form

$$S = \{x \mid Ax \preceq b, Fx = g\}.$$

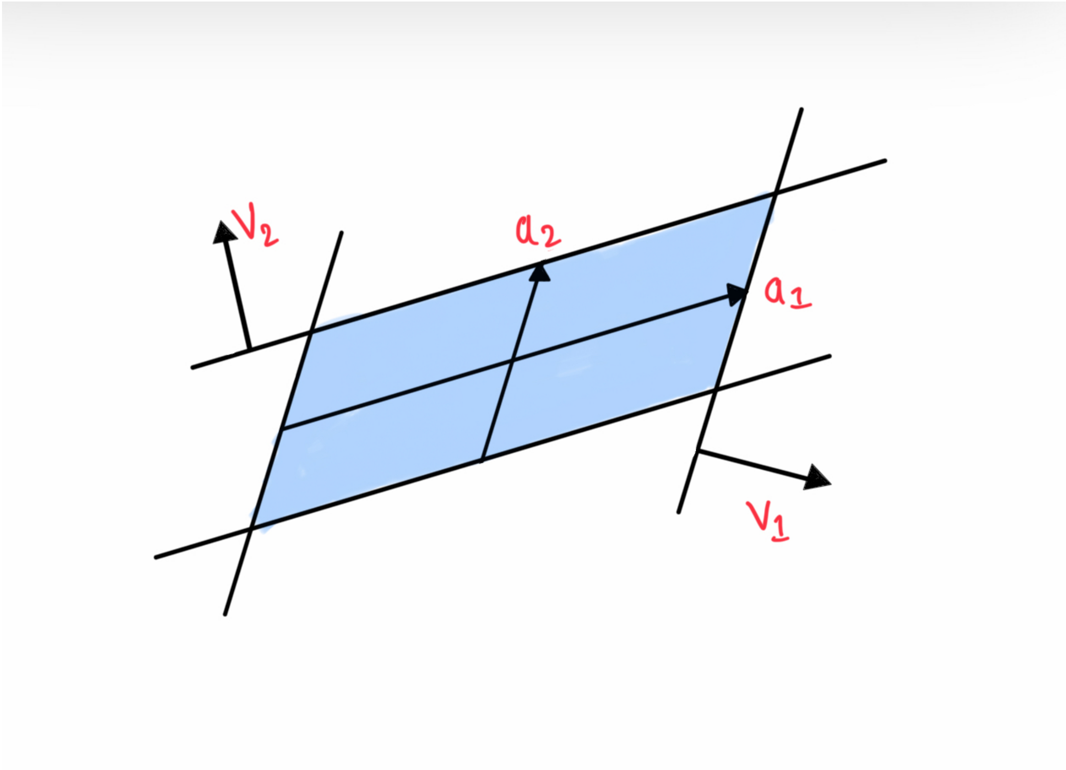
(a) $S = \{y_1 a_1 + y_2 a_2 \mid -1 \le y_1 \le 1, -1 \le y_2 \le 1\}$, where $a_1, a_2 \in \mathbb{R}^n$.
(b) $S = \{x \in \mathbb{R}^n \mid x \succeq 0, x^T a_1 = b_1, x^T a_2 = b_2\}$, where $a_1, \ldots, a_n \in \mathbb{R}^n$ and $b_1, b_2 \in \mathbb{R}$.
(c) $S = \{x \in \mathbb{R}^n \mid x \succeq 0, x^T y \le 1 \text{ for all } y \text{ with } \|y\|_2 = 1\}$.
(d) $S = \{x \in \mathbb{R}^n \mid x \succeq 0, x^T y \le 1 \text{ for all } y \text{ with } \sum_{i=1}^n |y_i| = 1\}$.

Solution:

(a) $S = \{y_1 a_1 + y_2 a_2 \mid -1 \le y_1 \le 1, -1 \le y_2 \le 1\}$, where $a_1, a_2 \in \mathbb{R}^n$.
S is a polyhedron. It is a parallelogram with corners:

$$a_1 + a_2,\ a_1 - a_2,\ -a_1 + a_2,\ -a_1 - a_2.$$

Let us assume that a_1 and a_2 are independent.
Polyhedron S can be considered as the intersection of three sets:



- S_1: The plane formed by a_1 and a_2.

$$S_2 = \{z + y_1 a_1 + y_2 a_2 \mid a_1^T z = 0,\ a_2^T z = 0,\ -1 \le y_1 \le 1\}$$

S_2 is a slab which is parallel to a_2 and orthogonal to S_1.

$$S_3 = \{z + y_1 a_1 + y_2 a_2 \mid a_1^T z = 0,\ a_2^T z = 0,\ -1 \le y_2 \le 1\}$$

S_3 is a slab which is parallel to a_1 and orthogonal to S_1.

The sets can be described as follows:

- S_1: It passes through the origin and is parallel to a_1, a_2.

  $$P_k^T x = 0 \quad k = 1, \dots, n-2$$

  P_k are $n-2$ independent vectors that are orthogonal to a_1 and a_2.

  Since they are orthogonal to a_1, a_2, we need not include them in the general equation. (That is why $n-2$).

- S_2: Since it is limited in the direction of a_1, we can say it includes all points between two hyperplanes with V_1 as the normal vector:

  $$V_1 = a_1 - \frac{a_1^T a_2}{\|a_2\|^2} a_2$$

  (General form of vector projection of a_1 perpendicular to a_2).

  The two hyperplanes will be of the form:

  $$H_1 : V_1^T z = d_1 \quad H_2 : V_1^T z = d_2$$

  Since S_2 is limited in the direction of a_1, H_1 has a_1 on it, and H_2 has $-a_1$ on it.

$$d_1 = V_1^T a_1, \quad d_2 = -V_1^T a_1$$

$$z \in S_2 \iff -|V_1^T a_1| \leq V_1^T z \leq |V_1^T a_1|$$

Now repeat this in the case of S_3:

Since S_3 is limited in the direction of a_2, it includes all points between hyperplanes with V_2 as the normal vector. One has a_2 on it and $-a_2$ on the other.

$$V_2 = a_2 - \frac{a_2^T a_1}{\|a_1\|^2} a_1$$

(General form of vector projection of a_2 perpendicular to a_1).

$$H_3 : V_2^T z = d_3, \quad H_4 : V_2^T z = d_4$$

$$d_3 = V_2^T a_2, \quad d_4 = -V_2^T a_2$$

$$z \in S_3 \iff -|V_2^T a_2| \leq V_2^T z \leq |V_2^T a_2|$$

Thus, $S = S_1 \cap S_2 \cap S_3$ with these equalities and inequalities:

$$\begin{gathered} P_k^T z = 0 \quad \forall k = 1, \ldots, n-2 \\ V_1^T z \leq |V_1^T a_1|, \quad -V_1^T z \leq |V_1^T a_1| \\ V_2^T z \leq |V_2^T a_2|, \quad -V_2^T z \leq |V_2^T a_2| \\ \therefore S \text{ is a polyhedron.} \end{gathered}$$

(b)

$$S = \{x \in \mathbb{R}^n \mid x \succeq 0, \, \mathbf{1}^T x = 1, \, \sum_{i=1}^n x_i a_i = b_1, \, \sum_{i=1}^n x_i a_i^2 = b_2\},$$

where $a_1, \ldots, a_n \in \mathbb{R}$ and $b_1, b_2 \in \mathbb{R}$.

$$x_k \succeq 0 \implies \text{Half-space bounded by hyperplane } x_k = 0.$$

$$\mathbf{1}^T x = 1 \implies \text{Hyperplane.}$$

$$\sum_{i=1}^n x_i a_i = b_1 \implies \mathbf{A}_1^T x = b_1 \implies \text{Hyperplane.}$$

$$\sum_{i=1}^n x_i a_i^2 = b_2 \implies \mathbf{A}_2^T x = b_2 \implies \text{Hyperplane,}$$

$$\mathbf{A}_2^T = [a_1^2 \, a_2^2 \, \ldots \, a_n^2].$$

S is an intersection of 3 hyperplanes and one half-space. By definition, S is a polyhedron.

(c)

$$S = \{x \in \mathbb{R}^n \mid x \succeq 0, \, x^T y \leq 1 \, \forall y \text{ with } \|y\|_2 = 1\}.$$

- $x \succeq 0$: Each component of x is non-negative ($x_i \geq 0$).
- $x^T y \leq 1$: For all y on the unit sphere ($\|y\|_2 = 1$), $x^T y$ must not exceed 1.

  Using the Cauchy-Schwarz inequality: For $x, y \in \mathbb{R}^n$,

$$x^T y \leq \|x\|_2 \|y\|_2.$$

  If $\|y\|_2 = 1$, then

$$x^T y \leq \|x\|_2.$$

  Given $x^T y \leq 1 \; \forall y$ with $\|y\|_2 = 1$,
Since $x^T y$ will be maximum when y is aligned with x, choose $y = \frac{x}{\|x\|_2}$.

$$x^T \frac{x}{\|x\|_2} \leq 1 \implies \|x\|_2 \leq 1.$$

  Using these results,

$$S = \{x \in \mathbb{R}^n \mid \|x\|_2 \leq 1, \, x \succeq 0\}.$$

  S is the intersection of:

- The unit ball: $\{x \in \mathbb{R}^n \mid \|x\|_2 \leq 1\}$, which is a non-polyhedral convex set.
- The non-negative orthant: $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0 \; \forall i\}$, which is a polyhedron.

Since a polyhedron is defined as a finite number of half-spaces, but the unit ball $\|x\|_2 \leq 1$ cannot be expressed as the intersection of a finite number of half-spaces because of its curved boundary which requires infinitely many half-spaces to represent.

- As S requires infinitely many half-spaces to represent, by definition, S is not a polyhedron.

(d)

$$S = \{x \in \mathbb{R}^n \mid x \succeq 0, \, x^T y \leq 1 \; \forall y \text{ with } \sum_{i=1}^{n} |y_i| = 1\}.$$

Let us consider these cases:
1. If $|x_i| \leq 1 \; \forall i$:
For all y with $\sum_{i=1}^{n} |y_i| = 1$, we have:

$$x^T y = \sum_{i=1}^{n} x_i y_i.$$

By using the triangle inequality:

$$\sum_{i=1}^{n} x_i y_i \leq \sum_{i=1}^{n} |x_i||y_i|.$$

Since $|x_i| \leq 1$ and $\sum_{i=1}^{n} |y_i| = 1$, we can say:

$$\sum_{i=1}^{n} |x_i||y_i| \leq \sum_{i=1}^{n} |y_i| \leq 1.$$

So, if $|x| \leq 1 \; \forall i$, then $x^T y \leq 1 \; \forall \; y$

2. If $x^T y \leq 1 \; \forall y$:
Let t be the index where $|x_t| = \max |x_i|$.

We need to choose y such that

$$y_t = 1, \quad \text{if } x_t > 0,$$

$$y_t = -1, \quad \text{if } x_t < 0,$$

$$y_i = 0, \quad \forall i \neq t.$$

Putting this y in the given inequality:

$$x^T y = \sum_i x_i y_i = x_t y_t = |x_t| = \max |x_i|.$$

Since $x^T y \leq 1 \ \forall y$,

$$\max |x_i| \leq 1.$$

Thus, $|x_i| \leq 1 \ \forall i$.
So we can say,

$$x^T y \leq 1 \quad \forall y \text{ with } \sum_{i=1}^{n} |y_i| = 1 \iff |x_i| \leq 1 \quad \forall i.$$

S can be rewritten as

$$S_0 = \{x \in \mathbb{R}^n : 0 \leq x_i \leq 1 \ \forall i = 1, \ldots, n\},$$

and it is a system of $2n$ linear equalities,

$$-x_i \leq 0, \quad x_i \leq 1 \ \forall i = 1, \ldots, n.$$

S_0 is the intersection of the non-negative orthant $\mathbb{R}^n_+$ and the set

$$\{x \in \mathbb{R}^n : |x_i| \leq 1 \ \forall i\}.$$

Since S is described as a finite number of linear inequalities, It is a polyhedron.