IIT Hyderabad **Scribed By: Rushil Kaul (2021101063)**
- **Siddharth Mangipudi (2021101060)**

Optimization Methods (CS1.404) **Lecture #4**
Instructor: Dr. Naresh Manwani **27^{th} Jan 2024**

Outline. *The lecture covers fundamental concepts in the Optimization Methods course, including affine and convex sets, their properties, and operations such as intersections and unions. It also discusses hyperplanes, half-spaces, and important theorems like Weierstrass' Theorem and the Closest Point Theorem. The main topics are:*

- ***Affine Sets***
  - *Definition and properties*
- ***Convex Sets and Convex Combinations***
  - *Definition of convex sets*
  - *Convex combinations*
  - *Intersection and union of convex sets*
- ***Hyperplanes and Half-spaces***
  - *Definition and equations*
- ***Key Theorems***
  - *Weierstrass' Theorem*
  - *Closest Point Theorem*

1 Affine Sets

Definition 1. A set $C \subseteq \mathbb{R}^d$ is called **affine** if, for any two distinct points in C, the entire affine line passing through these points also lies in C. Formally,

$$\text{If } \bar{x}_1, \bar{x}_2 \in C, \text{ then } \theta\bar{x}_1 + (1-\theta)\bar{x}_2 \in C, \quad \forall\theta \in \mathbb{R}.$$

Key properties of affine sets:

- A set is affine if and only if it contains every **affine combination** of its points.
- An affine set must include the **entire line** extending through any two of its points.

2 Convex Sets and Convex Combinations

2.1 Convex Sets

Definition 2. A set $X \subseteq \mathbb{R}^d$ is called **convex** if, for any two points in X, the line segment joining them also lies entirely within X. Formally,

$$\text{If } \bar{x}_1, \bar{x}_2 \in X, \text{ then } \lambda \bar{x}_1 + (1-\lambda)\bar{x}_2 \in X, \quad \forall \lambda \in [0,1].$$

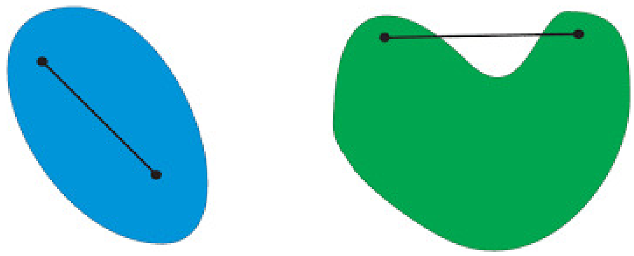

Figure 1: An example of a convex set (left), example of non-convex set (right).

Key properties of convex sets:

- Convexity ensures that only the **line segment** between $\bar{x}_1$ and $\bar{x}_2$ is included in X, not the entire line.
- Every **affine set** is convex, but the converse is not necessarily true; there exist convex sets that are not affine.

2.2 Convex Combinations

Definition 3. A **convex combination** of a set of points $\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_n$ is a **linear combination** of these points where all coefficients are **non-negative** and sum to 1. That is,

$$\bar{x} = \sum_{i=0}^{n} \lambda_i \bar{x}_i, \quad \text{where } \lambda_i \geq 0 \text{ and } \sum_{i=0}^{n} \lambda_i = 1.$$

This ensures that the resulting point $\bar{x}$ is a **weighted average** of the given points, staying within their convex hull.

2.3 Intersection of Convex Sets

Theorem 1. *The intersection of convex sets is convex. Specifically, if $X_1, X_2, \ldots, X_k$ are convex subsets of $\mathbb{R}^d$, then their intersection*

$$X = \bigcap_{i=1}^{k} X_i$$

is also convex.

Proof. Let $\bar{z}_1, \bar{z}_2 \in X$. By the definition of intersection, this means

$$\bar{z}_1, \bar{z}_2 \in X_i, \quad \forall i = 1, \ldots, k.$$

Since each X_i is convex, for any $\lambda \in [0, 1]$,

$$\lambda \bar{z}_1 + (1 - \lambda)\bar{z}_2 \in X_i, \quad \forall i = 1, \ldots, k.$$

Since this holds for all i, we conclude that

$$\lambda \bar{z}_1 + (1 - \lambda)\bar{z}_2 \in \bigcap_{i=1}^{k} X_i = X.$$

Thus, the intersection of convex sets remains convex. □ □

2.4 Union of Convex Sets is Not Convex

Theorem 2. *The union of convex sets is not necessarily convex.*

Proof. Consider the two convex sets C_1 and C_2 defined as:

$$C_1 = \{(\bar{x}_1, \bar{x}_2) \in \mathbb{R}^2 : \bar{x}_1^2 + \bar{x}_2^2 \leq 1\},$$

the unit disk, and

$$C_2 = \{(\bar{x}_1, \bar{x}_2) \in \mathbb{R}^2 : (\bar{x}_1 - 2)^2 + \bar{x}_2^2 \leq 1\},$$

a disk centered at $(2, 0)$.
Both C_1 and C_2 are convex, but their union $C_1 \cup C_2$ is not convex. For example, the line segment joining the points $(0, 0)$ in C_1 and $(2, 0)$ in C_2 passes outside of the union, violating the convexity condition.
Thus, the union of convex sets is not necessarily convex. □ □

2.5 Hyperplane and Half-space

A **hyperplane** in $\mathbb{R}^d$ is the set of points $\bar{x}$ that satisfy the equation:

$$\{\bar{x} \mid \bar{w}^T \bar{x} = b\},$$

where $\bar{w} \in \mathbb{R}^d$ is a vector normal to the hyperplane, and $b \in \mathbb{R}$ is the offset.
- The vector $\bar{w}$ defines the orientation of the hyperplane. - The scalar b controls the position of the hyperplane relative to the origin.
A **half-space** is one of the two parts of $\mathbb{R}^d$ divided by a hyperplane. A half-space can be represented by one of the following sets:

$$\{\bar{x} \mid \bar{w}^T \bar{x} \leq b\} \quad \text{or} \quad \{\bar{x} \mid \bar{w}^T \bar{x} \geq b\}.$$

- A half-space is **convex** because, for any two points in the half-space, the line segment joining them remains inside the half-space. - We distinguish between **closed** and **open** half-spaces: - **Closed half-space**: $\{\bar{x} \mid \bar{w}^T \bar{x} \leq b\}$ includes the hyperplane. - **Open half-space**: $\{\bar{x} \mid \bar{w}^T \bar{x} < b\}$ excludes the hyperplane.
Both open and closed half-spaces are convex.

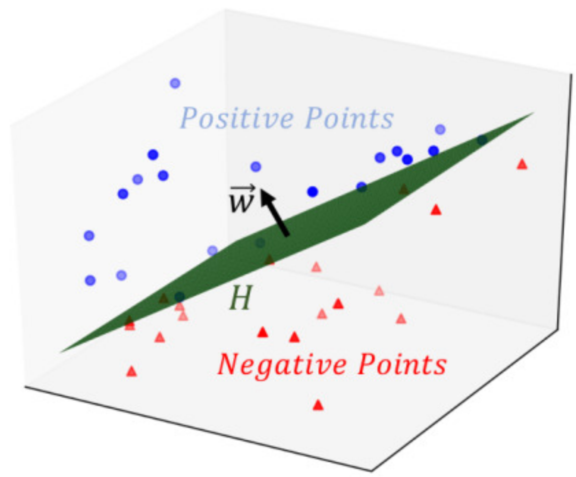

Figure 2: An example of a hyperplane in $\mathbb{R}^3$.

3 Weierstrass Theorem

Theorem 3 (Weierstrass Theorem). *Let $X \subseteq \mathbb{R}^n$ be a non-empty compact (closed and bounded) set, and let $f : X \to \mathbb{R}$ be a continuous function on X. Then, f attains both a minimum and a maximum on X, i.e., there exist points $\bar{x}_{min}, \bar{x}_{max} \in X$ such that:*

$$f(\bar{x}_{min}) \leq f(\bar{x}) \leq f(\bar{x}_{max}), \quad \forall \bar{x} \in X.$$

Proof. Since $X \subseteq \mathbb{R}^n$ is compact (closed and bounded) and $f : X \to \mathbb{R}$ is continuous, the image $f(X)$ is also compact in $\mathbb{R}$. By the Heine-Borel theorem, $f(X)$ is closed and bounded. Because $f(X)$ is bounded, the supremum $M = \sup_{\bar{x}\in X} f(\bar{x})$ and infimum $m = \inf_{\bar{x}\in X} f(\bar{x})$ exist in $\mathbb{R}$. Moreover, since $f(X)$ is closed, it contains its supremum and infimum, meaning $M, m \in f(X)$.
To show that f attains its maximum, consider a sequence $\{\bar{x}_k\} \subseteq X$ such that $f(\bar{x}_k) \to M$. By the Bolzano-Weierstrass theorem, there exists a convergent subsequence $\{\bar{x}_{k_j}\}$ that converges to some $\bar{x}_{\max} \in X$ (as X is closed).
By the continuity of f, it follows that $f(\bar{x}_{\max}) = \lim_{j\to\infty} f(\bar{x}_{k_j}) = M$. Similarly, for the minimum, take a sequence $\{\bar{y}_k\} \subseteq X$ with $f(\bar{y}_k) \to m$. Extract a convergent subsequence $\{\bar{y}_{k_j}\}$ converging to $\bar{x}_{\min} \in X$. By continuity, $f(\bar{x}_{\min}) = \lim_{j\to\infty} f(\bar{y}_{k_j}) = m$.
Therefore, f attains its maximum and minimum values at $\bar{x}_{\max}$ and $\bar{x}_{\min}$, respectively, completing the proof. □

4 Closest Point Theorem

Theorem 4 (Closest Point Theorem). *Let $S \subseteq \mathbb{R}^n$ be a non-empty closed convex set, and let $\bar{y} \notin S$. Then, there exists a unique point $\bar{x}_0 \in S$ such that the distance from $\bar{y}$ to $\bar{x}_0$ is minimized. Specifically, the point $\bar{x}_0$ minimizes the distance between $\bar{y}$ and points in S, i.e.,*

$$dist(\bar{y}, \bar{x}_0) = \min_{\bar{x}\in S} dist(\bar{y}, \bar{x}).$$

Furthermore, $\bar{x}_0$ is the closest point to $\bar{y}$ if and only if the following condition holds:

$$(\bar{y} - \bar{x}_0)^T(\bar{x} - \bar{x}_0) \leq 0 \quad \forall \bar{x} \in S.$$

This condition implies that the vector $\bar{y} - \bar{x}_0$ forms an angle of at least 90 degrees with any vector $\bar{x} - \bar{x}_0$ for $\bar{x} \in S$.

Proof. Let $f(\bar{x}) = \|\bar{x} - \bar{y}\|^2$. We want to find the minimizer of $f(\bar{x})$ over S. Formally, we are looking for:

$$\min_{\bar{x} \in S} f(\bar{x}) = \min_{\bar{x} \in S} \|\bar{x} - \bar{y}\|^2.$$

Since $f(\bar{x})$ is continuous over S, and S is closed, we can apply the Weierstrass Theorem. However, S may not be bounded, so the minimum might not exist on S alone.

To resolve this, consider a point $\bar{z} \in S$ and define $r = \|\bar{z} - \bar{y}\|$. Let $S_1 = S \cap B[\bar{y}, r]$, where $B[\bar{y}, r]$ is the closed ball centered at $\bar{y}$ with radius r. The set S_1 is closed and bounded, so by the Weierstrass Theorem, a minimum of $f(\bar{x})$ exists on S_1.

Let $\bar{x}_0 = \arg\min_{\bar{x} \in S_1} \|\bar{x} - \bar{y}\|^2$. It is easy to verify that:

$$\|\bar{y} - \bar{x}_0\| < \min_{\bar{x} \in S \setminus S_1} \|\bar{x} - \bar{y}\|^2,$$

showing that $\bar{x}_0$ is indeed the closest point to $\bar{y}$.

Now, to show the uniqueness of $\bar{x}_0$, suppose there exists another point $\bar{x}_1 \in S$ such that $\|\bar{y} - \bar{x}_1\| = \|\bar{y} - \bar{x}_0\| = \gamma$. Consider the point

$$\bar{x}_2 = \frac{\bar{x}_0 + \bar{x}_1}{2}.$$

We now compute the distance $\|\bar{y} - \bar{x}_2\|$:

$$\|\bar{y} - \bar{x}_2\| = \left\| \bar{y} - \frac{\bar{x}_0 + \bar{x}_1}{2} \right\| = \left\| \frac{1}{2}(\bar{y} - \bar{x}_0) + \frac{1}{2}(\bar{y} - \bar{x}_1) \right\|.$$

By the triangle inequality, we have:

$$\|\bar{y} - \bar{x}_2\| \leq \frac{1}{2}\|\bar{y} - \bar{x}_0\| + \frac{1}{2}\|\bar{y} - \bar{x}_1\| = \frac{1}{2}\gamma + \frac{1}{2}\gamma = \gamma.$$

Case 1: Strict Inequality If $\|\bar{y} - \bar{x}_2\| < \gamma$, this would imply $\bar{x}_2$ is closer to $\bar{y}$ than both $\bar{x}_0$ and $\bar{x}_1$, contradicting minimality.

Case 2: Equality Condition Equality $\|\bar{y} - \bar{x}_2\| = \gamma$ holds if and only if $\bar{y} - \bar{x}_1 = k(\bar{y} - \bar{x}_0)$ for some $k \geq 0$. If $\bar{x}_0 \neq \bar{x}_1$, this requires collinearity of $\bar{y}, \bar{x}_0, \bar{x}_1$. Substituting $k = 1$ yields $\bar{y} - \bar{x}_1 = \bar{y} - \bar{x}_0$, forcing $\bar{x}_0 = \bar{x}_1$. Thus, equality cannot occur for distinct $\bar{x}_0, \bar{x}_1$.

The strict convexity of $f(\bar{x})$ on convex S provides an alternative proof that the inequality must be strict. Therefore, $\bar{x}_0$ is unique.

Part a: Proof of Closest Point Condition

We now show that $\bar{x}_0$ is the closest point to $\bar{y}$ if and only if:

$$(\bar{y} - \bar{x}_0)^T(\bar{x} - \bar{x}_0) \leq 0 \quad \forall \bar{x} \in S.$$

Let $\bar{x}_0$ be the point such that:

$$(\bar{y} - \bar{x}_0)^T(\bar{x} - \bar{x}_0) \leq 0 \quad \forall \bar{x} \in S.$$

We need to show that $\bar{x}_0$ is the closest point to $\bar{y}$. First, let's expand $\|\bar{y} - \bar{x}\|^2$ for any point $\bar{x} \in S$:

$$\|\bar{y} - \bar{x}\|^2 = \|(\bar{y} - \bar{x}_0) + (\bar{x}_0 - \bar{x})\|^2.$$

Expanding the square:

$$\|\bar{y} - \bar{x}\|^2 = \|\bar{y} - \bar{x}_0\|^2 + \|\bar{x}_0 - \bar{x}\|^2 + 2(\bar{y} - \bar{x}_0)^T(\bar{x}_0 - \bar{x}).$$

By the given condition, we know that:

$$(\bar{y} - \bar{x}_0)^T(\bar{x}_0 - \bar{x}) \leq 0.$$

This implies:

$$\|\bar{y} - \bar{x}\|^2 \geq \|\bar{y} - \bar{x}_0\|^2 + \|\bar{x}_0 - \bar{x}\|^2.$$

Thus, we conclude that:

$$\|\bar{y} - \bar{x}\|^2 \geq \|\bar{y} - \bar{x}_0\|^2 \quad \forall \bar{x} \in S.$$

Since the squared distance is minimized when $\bar{x} = \bar{x}_0$, we conclude that $\bar{x}_0$ is the closest point to $\bar{y}$. Thus, $\bar{x}_0$ is the unique point minimizing the distance from $\bar{y}$ to S. $\square$

Part b:

Let $\bar{x}_0$ be the closest point to $\bar{y}$, i.e., the point $\bar{x}_0 \in S$ such that:

$$\|\bar{y} - \bar{x}_0\|^2 \leq \|\bar{y} - \bar{x}\|^2 \quad \forall \bar{x} \in S.$$

Now, consider a point $\bar{x}_1 \in S$, and let $\bar{x}_2 = (1 - \lambda)\bar{x}_0 + \lambda\bar{x}_1$, where $\lambda \in (0, 1]$ which ensures $\bar{x}_2$ lies on the line segment between $\bar{x}_0$ and $\bar{x}_1$. We want to show that:

$$\|\bar{y} - \bar{x}_2\|^2 \geq \|\bar{y} - \bar{x}_0\|^2.$$

First, expand $\|\bar{y} - \bar{x}_2\|^2$:

$$\|\bar{y} - \bar{x}_2\|^2 = \|\bar{y} - (1 - \lambda)\bar{x}_0 - \lambda\bar{x}_1\|^2.$$

This simplifies to:

$$\|\bar{y} - \bar{x}_2\|^2 = \|(\bar{y} - \bar{x}_0) - \lambda(\bar{x}_1 - \bar{x}_0)\|^2.$$

Expanding the square:

$$\|\bar{y} - \bar{x}_2\|^2 = \|\bar{y} - \bar{x}_0\|^2 + \lambda^2\|\bar{x}_1 - \bar{x}_0\|^2 - 2\lambda(\bar{y} - \bar{x}_0)^T(\bar{x}_1 - \bar{x}_0).$$

Now, we know that $\|\bar{y} - \bar{x}_0\|^2 \leq \|\bar{y} - \bar{x}\|^2$ for all $\bar{x} \in S$, so:

$$\lambda^2\|\bar{x}_1 - \bar{x}_0\|^2 - 2\lambda(\bar{y} - \bar{x}_0)^T(\bar{x}_1 - \bar{x}_0) \geq 0.$$

Taking the limit as $\lambda \to 0^+$, we get:

$$-2(\bar{y} - \bar{x}_0)^T(\bar{x}_1 - \bar{x}_0) \geq 0,$$

which implies:

$$(\bar{y} - \bar{x}_0)^T(\bar{x}_1 - \bar{x}_0) \leq 0.$$

Thus, for any point $\bar{x}_1 \in S$, we have:

$$(\bar{y} - \bar{x}_0)^T(\bar{x}_1 - \bar{x}_0) \leq 0.$$

This completes the proof of part b. $\square$

$\square$

References

[1] Convex Optimization by Stephen Boyd

[2] An Introduction to Optimization by Chong and Zak

[3] Introduction to Nonlinear Optimization by Amir Beck