MS-E2122 - Nonlinear Optimization
Lecture II

Fernando Dias

Department of Mathematics and Systems Analysis

Aalto University
School of Science

September 3, 2023

Outline of this lecture

Last Week...

- Introduction: optimisation, variables, constraints, objective function
- Example of Applications

Convex sets and optimisation

"...in fact, the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity."
R. Tyrrell Rockafellar, in SIAM Review, 1993

(N. Maculan with T. Rockafellar at EURO 2015)

Outline of this lecture

Convex sets

Convex set: a set that contains all elements of the line segment connecting any two of its elements.

Let $S \subseteq \mathbb{R}^n$ and $x_j \in S$ for $j = 1, \dots, k$. Some key concepts:

- **Linear combination:** $\sum_{j=1}^{k} \lambda_j x_j$, where $\lambda_j \in \mathbb{R}$ for $j = 1, \dots, k$;

Convex sets

Convex set: a set that contains all elements of the line segment connecting any two of its elements.

Let $S \subseteq \mathbb{R}^n$ and $x_j \in S$ for $j = 1, \dots, k$. Some key concepts:

- **Linear combination:** $\sum_{j=1}^{k} \lambda_j x_j$, where $\lambda_j \in \mathbb{R}$ for $j = 1, \dots, k$;
- **Affine combination:** $\sum_{j=1}^{k} \lambda_j x_j$, where $\lambda_j \in \mathbb{R}$ for $j = 1, \dots, k$ and $\sum_{j=1}^{k} \lambda_j = 1$;

Convex sets

Convex set: a set that contains all elements of the line segment connecting any two of its elements.

Let $S \subseteq \mathbb{R}^n$ and $x_j \in S$ for $j = 1, \dots, k$. Some key concepts:

- **Linear combination:** $\sum_{j=1}^{k} \lambda_j x_j$, where $\lambda_j \in \mathbb{R}$ for $j = 1, \dots, k$;
- **Affine combination:** $\sum_{j=1}^{k} \lambda_j x_j$, where $\lambda_j \in \mathbb{R}$ for $j = 1, \dots, k$ and $\sum_{j=1}^{k} \lambda_j = 1$;
- **Conic combination:** $\sum_{j=1}^{k} \lambda_j x_j$, where $\lambda_j \geq 0$ for $j = 1, \dots, k$;

Convex sets

Convex set: a set that contains all elements of the line segment connecting any two of its elements.

Let $S \subseteq \mathbb{R}^n$ and $x_j \in S$ for $j = 1, \dots, k$. Some key concepts:

- **Linear combination:** $\sum_{j=1}^{k} \lambda_j x_j$, where $\lambda_j \in \mathbb{R}$ for $j = 1, \dots, k$;
- **Affine combination:** $\sum_{j=1}^{k} \lambda_j x_j$, where $\lambda_j \in \mathbb{R}$ for $j = 1, \dots, k$ and $\sum_{j=1}^{k} \lambda_j = 1$;
- **Conic combination:** $\sum_{j=1}^{k} \lambda_j x_j$, where $\lambda_j \geq 0$ for $j = 1, \dots, k$;
- **Convex combination:** $\sum_{j=1}^{k} \lambda_j x_j$, where $\lambda_j \geq 0$ for $j = 1, \dots, k$ and $\sum_{j=1}^{k} \lambda_j = 1$.

Convex sets

Definition 1 (Convex sets)

A **set** $S \subseteq \mathbb{R}^n$ is said to be **convex** if $\overline{x} = \sum_{j=1}^{k} \lambda_j x_j$ belongs to S, where $\sum_{j=1}^{k} \lambda_j = 1$, $\lambda_j \geq 0$ and $x_j \in S$ for $j = 1, \ldots, k$.

Convex sets

Definition 1 (Convex sets)

A **set** $S \subseteq \mathbb{R}^n$ is said to be **convex** if $\overline{x} = \sum_{j=1}^{k} \lambda_j x_j$ belongs to S, where $\sum_{j=1}^{k} \lambda_j = 1$, $\lambda_j \geq 0$ and $x_j \in S$ for $j = 1, \dots, k$.

Convexity preserving operations as a consequence of Definition 1:

Lemma 2

Let S_1 and S_2 be **convex sets** *in $\mathbb{R}^n$. Then these are also* **convex***:*

1. **Intersection:** $S_1 \cap S_2$;

Convex sets

Definition 1 (Convex sets)

A **set** $S \subseteq \mathbb{R}^n$ is said to be **convex** if $\overline{x} = \sum_{j=1}^{k} \lambda_j x_j$ belongs to S, where $\sum_{j=1}^{k} \lambda_j = 1$, $\lambda_j \geq 0$ and $x_j \in S$ for $j = 1, \dots, k$.

Convexity preserving operations as a consequence of Definition 1:

Lemma 2

Let S_1 and S_2 be **convex sets** *in $\mathbb{R}^n$. Then these are also* **convex***:*

1. **Intersection:** $S_1 \cap S_2$;
2. **Minkowski addition:** $S_1 + S_2 = \{x_1 + x_2 : x_1 \in S_1, x_2 \in S_2\}$;

Convex sets

Definition 1 (Convex sets)

A **set** $S \subseteq \mathbb{R}^n$ is said to be **convex** if $\overline{x} = \sum_{j=1}^{k} \lambda_j x_j$ belongs to S, where $\sum_{j=1}^{k} \lambda_j = 1$, $\lambda_j \geq 0$ and $x_j \in S$ for $j = 1, \dots, k$.

Convexity preserving operations as a consequence of Definition 1:

Lemma 2

Let S_1 and S_2 be **convex sets** *in $\mathbb{R}^n$. Then these are also* **convex***:*

1. **Intersection:** $S_1 \cap S_2$;
2. **Minkowski addition:** $S_1 + S_2 = \{x_1 + x_2 : x_1 \in S_1, x_2 \in S_2\}$;
3. **Minkowski difference:** $S_1 - S_2 = \{x_1 - x_2 : x_1 \in S_1, x_2 \in S_2\}$;

Convex sets

Definition 1 (Convex sets)

A **set** $S \subseteq \mathbb{R}^n$ is said to be **convex** if $\overline{x} = \sum_{j=1}^{k} \lambda_j x_j$ belongs to S, where $\sum_{j=1}^{k} \lambda_j = 1$, $\lambda_j \geq 0$ and $x_j \in S$ for $j = 1, \dots, k$.

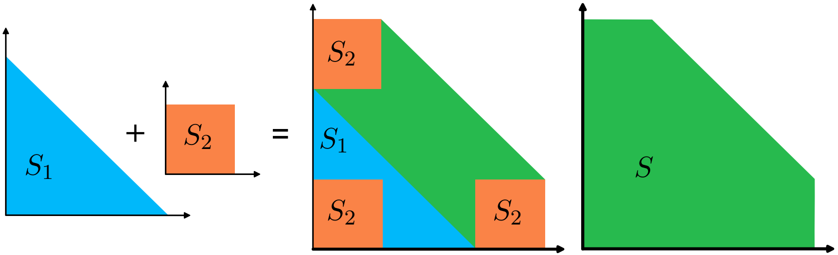
Convexity preserving operations as a consequence of Definition 1:

Lemma 2

Let S_1 and S_2 be **convex sets** *in $\mathbb{R}^n$. Then these are also* **convex***:*

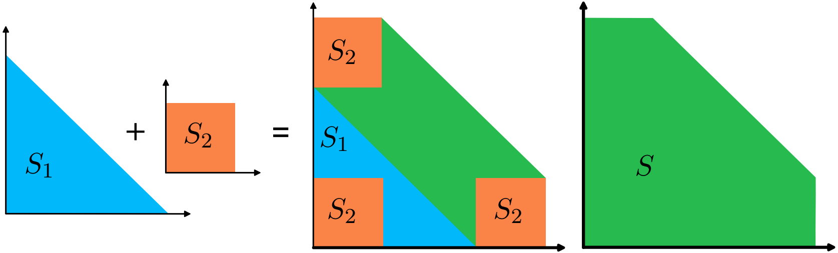
1. **Intersection:** $S_1 \cap S_2$;
2. **Minkowski addition:** $S_1 + S_2 = \{x_1 + x_2 : x_1 \in S_1, x_2 \in S_2\}$;
3. **Minkowski difference:** $S_1 - S_2 = \{x_1 - x_2 : x_1 \in S_1, x_2 \in S_2\}$;
4. **Linear transformation:** $S = \{Ax_1 + b : x_1 \in S_1\}$.

Convexity preserving operations

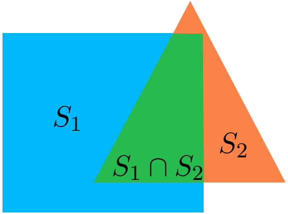


Minkowski sum of two convex sets.

Convexity preserving operations



Minkowski sum of two convex sets.



Intersection of two convex sets.

Convex sets

Some examples of convex sets:

- **half-spaces**: $S = \{x : p^\top x \leq \alpha\} \subset \mathbb{R}^n$;

Convex sets

Some examples of convex sets:

- **half-spaces**: $S = \{x : p^\top x \leq \alpha\} \subset \mathbb{R}^n$;
- **hyperplanes**: $S = \{x : p^\top x = \alpha\} \subset \mathbb{R}^n$, where $p \neq 0^n$ is the normal vector and $\alpha \in \mathbb{R}$ is a scalar. S can be equivalently represented as $S = \{x : p^\top (x - \overline{x}) = 0\}$ for $\overline{x} \in S$;

Convex sets

Some examples of convex sets:

- **half-spaces**: $S = \left\{x : p^\top x \leq \alpha\right\} \subset \mathbb{R}^n$;
- **hyperplanes**: $S = \left\{x : p^\top x = \alpha\right\} \subset \mathbb{R}^n$, where $p \neq 0^n$ is the normal vector and $\alpha \in \mathbb{R}$ is a scalar. S can be equivalently represented as $S = \left\{x : p^\top (x - \overline{x}) = 0\right\}$ for $\overline{x} \in S$;
- **polyhedral sets**: $S = \{x : Ax \leq b\} \subset \mathbb{R}^n$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$;

Convex sets

Some examples of convex sets:

- **half-spaces**: $S = \left\{x : p^\top x \leq \alpha\right\} \subset \mathbb{R}^n$;
- **hyperplanes**: $S = \left\{x : p^\top x = \alpha\right\} \subset \mathbb{R}^n$, where $p \neq 0^n$ is the normal vector and $\alpha \in \mathbb{R}$ is a scalar. S can be equivalently represented as $S = \left\{x : p^\top (x - \overline{x}) = 0\right\}$ for $\overline{x} \in S$;
- **polyhedral sets**: $S = \{x : Ax \leq b\} \subset \mathbb{R}^n$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$;
- **norm-induced sets (balls):** $S = \{x : ||x - \overline{x}|| \leq \alpha\}$, where $||\cdot||$ is any norm and α a scalar;

Convex sets

Some examples of convex sets:

- **half-spaces**: $S = \{x : p^\top x \leq \alpha\} \subset \mathbb{R}^n$;
- **hyperplanes**: $S = \{x : p^\top x = \alpha\} \subset \mathbb{R}^n$, where $p \neq 0^n$ is the normal vector and $\alpha \in \mathbb{R}$ is a scalar. S can be equivalently represented as $S = \{x : p^\top (x - \overline{x}) = 0\}$ for $\overline{x} \in S$;
- **polyhedral sets**: $S = \{x : Ax \leq b\} \subset \mathbb{R}^n$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$;
- **norm-induced sets (balls):** $S = \{x : ||x - \overline{x}|| \leq \alpha\}$, where $|| \cdot ||$ is any norm and α a scalar;
- **norm cones**: $S = \{(x, \alpha) : ||x|| \leq \alpha\}$;

Convex sets

Some examples of convex sets:

- **half-spaces**: $S = \{x : p^\top x \leq \alpha\} \subset \mathbb{R}^n$;
- **hyperplanes**: $S = \{x : p^\top x = \alpha\} \subset \mathbb{R}^n$, where $p \neq 0^n$ is the normal vector and $\alpha \in \mathbb{R}$ is a scalar. S can be equivalently represented as $S = \{x : p^\top (x - \overline{x}) = 0\}$ for $\overline{x} \in S$;
- **polyhedral sets**: $S = \{x : Ax \leq b\} \subset \mathbb{R}^n$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$;
- **norm-induced sets (balls):** $S = \{x : ||x - \overline{x}|| \leq \alpha\}$, where $|| \cdot ||$ is any norm and α a scalar;
- **norm cones**: $S = \{(x, \alpha) : ||x|| \leq \alpha\}$;

Remark: p-norm: $||x||_p = \left(\sum_{j=1}^n |x_j|^p\right)^{\frac{1}{p}}$

Hyperplanes and halfspaces

A hyperplane $H = \{x : p^\top x = \alpha\}$, with normal $p \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, defines two closed half-spaces:

- $H^+ = \{x : p^\top x \geq \alpha\}$
- $H^- = \{x : p^\top x \leq \alpha\}$

Hyperplanes and halfspaces

A hyperplane $H = \{x : p^\top x = \alpha\}$, with normal $p \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, defines two closed half-spaces:

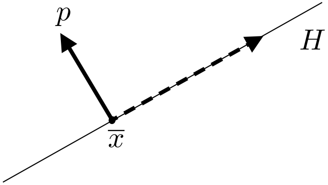
- $H^+ = \{x : p^\top x \geq \alpha\}$
- $H^- = \{x : p^\top x \leq \alpha\}$

Notice that H and its half-spaces can be defined in reference to any point $\overline{x} \in H$: $p^\top \overline{x} = \alpha$ must hold and thus

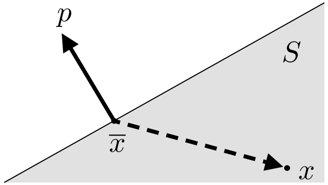
$$p^\top (x - \overline{x}) = p^\top \overline{x} - p^\top x = \alpha - \alpha = 0$$

can also be used to describe H. Also, $H^+ = \{x : p^\top (x - \overline{x}) \geq 0\}$ and $H^- = \{x : p^\top (x - \overline{x}) \leq 0\}$.

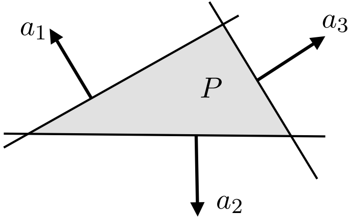
Hyperplanes and halfspaces



$H = \{x \in \mathbb{R}^n : p^\top(x - \overline{x}) = 0\}$ is a hyperplane with normal vector p displaced to $\overline{x}$.



$S = \{x \in \mathbb{R}^n : p^\top(x - \overline{x}) \leq 0\}$ is a halfspace defined by H.



A polyhedron P formed by the intersection of three halfspaces.

Convex hulls

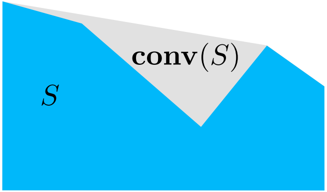
Definition 3 (Convex hull of a set)

Let $S \subseteq \mathbb{R}^n$ be an arbitrary set. The **convex hull** of S, denoted by $\mathbf{conv}(S)$, is the collection of all **convex combinations** of S. That is, for $x_j \in S$, $j = 1, \dots, k$, $x \in \mathbf{conv}(S)$ if and only if

$$x = \sum_{j=1}^{k} \lambda_j x_j$$

$$\sum_{j=1}^{k} \lambda_j = 1$$

$\lambda_j \geq 0$, for $j = 1, \dots, k$.



Example of an arbitrary set S and its convex hull $\mathbf{conv}(S)$.

Convex hulls

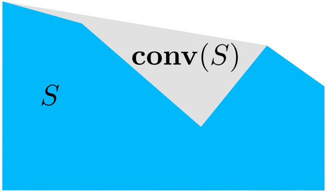
Definition 3 (Convex hull of a set)

Let $S \subseteq \mathbb{R}^n$ be an arbitrary set. The **convex hull** of S, denoted by $\mathbf{conv}(S)$, is the collection of all **convex combinations** of S. That is, for $x_j \in S$, $j = 1, \dots, k$, $x \in \mathbf{conv}(S)$ if and only if

$$x = \sum_{j=1}^{k} \lambda_j x_j$$

$$\sum_{j=1}^{k} \lambda_j = 1$$

$$\lambda_j \geq 0, \text{for } j = 1, \dots, k.$$



Example of an arbitrary set S and its convex hull $\mathbf{conv}(S)$.

Remark: from Definition 3 we notice that $\mathbf{conv}(S)$ is the tightest enveloping set that contains S. This can be shown by noticing that $\mathbf{conv}(S)$ is the intersection of all convex sets containing S.

Outline of this lecture

Closure and interior of a set

Let $N_\epsilon(x) = \{y : ||y - x|| < \epsilon\}$ denote an ϵ-neighbourhood of $x \in \mathbb{R}^n$, and let $S \subseteq \mathbb{R}^n$ be an arbitrary set.

Closure and interior of a set

Let $N_\epsilon(x) = \{y : ||y - x|| < \epsilon\}$ denote an ϵ-neighbourhood of $x \in \mathbb{R}^n$, and let $S \subseteq \mathbb{R}^n$ be an arbitrary set. We can now define the following key attributes of S:

1. **Interior of S:** $x \in \mathbf{int}(S)$ if $N_\epsilon(x) \subset S$ for *some* $\epsilon > 0$.
   S is solid if $\mathbf{int}(S) \neq \emptyset$ and open if $S = \mathbf{int}(S)$.

Closure and interior of a set

Let $N_\epsilon(x) = \{y : ||y - x|| < \epsilon\}$ denote an ϵ-neighbourhood of $x \in \mathbb{R}^n$, and let $S \subseteq \mathbb{R}^n$ be an arbitrary set. We can now define the following key attributes of S:

1. **Interior of S:** $x \in \mathbf{int}(S)$ if $N_\epsilon(x) \subset S$ for *some* $\epsilon > 0$.
   S is solid if $\mathbf{int}(S) \neq \emptyset$ and open if $S = \mathbf{int}(S)$.
2. **Boundary of S:** $x \in \mathbf{bou}(S)$ if $N_\epsilon(x)$ contains at least one point in S and one point not in S for *every* $\epsilon > 0$. Moreover, S is bounded if $S \subset N_\epsilon(x)$ for *some* $\epsilon > 0$, and S is compact if S is both closed and bounded.

Closure and interior of a set

Let $N_\epsilon(x) = \{y : ||y - x|| < \epsilon\}$ denote an ϵ-neighbourhood of $x \in \mathbb{R}^n$, and let $S \subseteq \mathbb{R}^n$ be an arbitrary set. We can now define the following key attributes of S:

1. **Interior of S:** $x \in \mathbf{int}(S)$ if $N_\epsilon(x) \subset S$ for *some* $\epsilon > 0$.
   S is solid if $\mathbf{int}(S) \neq \emptyset$ and open if $S = \mathbf{int}(S)$.
2. **Boundary of S:** $x \in \mathbf{bou}(S)$ if $N_\epsilon(x)$ contains at least one point in S and one point not in S for *every* $\epsilon > 0$. Moreover, S is bounded if $S \subset N_\epsilon(x)$ for *some* $\epsilon > 0$, and S is compact if S is both closed and bounded.
3. **Closure of S:** $x \in \mathbf{clo}(S)$ if $S \cap N_\epsilon(x) \neq \emptyset$ for *every* $\epsilon > 0$.
   S is closed is $S = \mathbf{clo}(S)$.

Closure and interior of a set

If S is a **convex set**, we can infer the **convexity** of $\mathbf{int}(S)$ and $\mathbf{clo}(S)$.

Theorem 4

Let $S \subseteq \mathbb{R}^n$ *be a convex set with* $\mathbf{int}(S) \neq \emptyset$. *Let* $x_1 \in \mathbf{clo}(S)$ *and* $x_2 \in \mathbf{int}(S)$. *Then* $x = \lambda x_1 + (1-\lambda)x_2 \in \mathbf{int}(S)$ *for all* $\lambda \in (0, 1)$.

Closure and interior of a set

If S is a **convex set**, we can infer the **convexity** of $\mathbf{int}(S)$ and $\mathbf{clo}(S)$.

Theorem 4

Let $S \subseteq \mathbb{R}^n$ be a convex set with $\mathbf{int}(S) \neq \emptyset$. Let $x_1 \in \mathbf{clo}(S)$ and $x_2 \in \mathbf{int}(S)$. Then $x = \lambda x_1 + (1-\lambda)x_2 \in \mathbf{int}(S)$ for all $\lambda \in (0,1)$.

Theorem 4 gives rise to the following corollaries:

Corollary 5

Let S be a convex set with $\mathbf{int}(S) \neq \emptyset$. Then

1. $\mathbf{int}(S)$ *is convex;*
2. $\mathbf{clo}(S)$ *is convex;*
3. $\mathbf{clo}(\mathbf{int}(S)) = \mathbf{clo}(S)$;
4. $\mathbf{int}(\mathbf{clo}(S)) = \mathbf{int}(S)$.

Weierstrass theorem

This **result** is used to guarantee the existence of optimal solutions (minimising/maximising) for optimisation problems. Let

$$(P) \; : \; z = \text{min. } \{f(x) : x \in S\}$$

be our optimisation problem. If an optimal solution x^* exists, then $f(x^*) \leq f(x)$ for all $x \in S$ and $z = f(x^*) = \min\{f(x) : x \in S\}$.

Weierstrass theorem

This **result** is used to guarantee the existence of optimal solutions (minimising/maximising) for optimisation problems. Let

$$(P) \ : \ z = \text{min. } \{f(x) : x \in S\}$$

be our optimisation problem. If an optimal solution x^* exists, then $f(x^*) \leq f(x)$ for all $x \in S$ and $z = f(x^*) = \min\{f(x) : x \in S\}$.

If an optimal solution cannot be attained, it might still be possible to obtain the infimum (or supremum for maximisation problems):

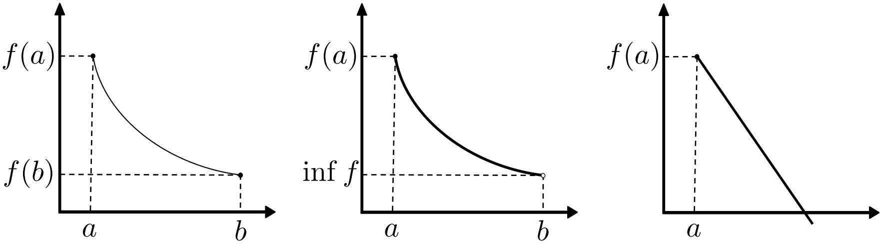
$$(P) : \ z = \inf\{f(x) : x \in S\}$$

which is the greatest lower bound of f in S. That is, $z \leq f(x)$ for all $x \in S$ and there is no $\overline{z} > z$ such that $\overline{z} \leq f(x)$ for all $x \in S$.

Remark: notice that $\min$ stands for minimum while min. stands for minimise.

Weierstrass theorem

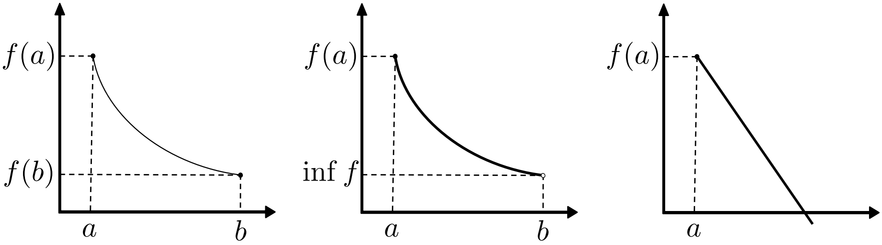
Weierstrass theorem describes when "inf" can be "safely" replaced by "min".



Examples of attainable minimum (left) and infimum (centre) and an example where neither are attainable (right).

Weierstrass theorem

Weierstrass theorem describes when "inf" can be "safely" replaced by "min".



Examples of attainable minimum (left) and infimum (centre) and an example where neither are attainable (right).

Theorem 6 (Weierstrass theorem)

Let $S \neq \emptyset$ be a compact set, and let $f : S \to \mathbb{R}$ be continuous on S. Then there is a minimising solution to $(P) :$ min. $\{f(x) : x \in S\}$.

Outline of this lecture

Separation and support of sets

Supporting hyperplanes and the separation of disjoint convex sets are key concepts for establishing optimality conditions and duality relationships.

Given a convex set S and a point $y \notin S$, we want to find $\overline{x} \in S$ closest to y and a hyperplane that separates them.

Separation and support of sets

Supporting hyperplanes and the separation of disjoint convex sets are key concepts for establishing optimality conditions and duality relationships.

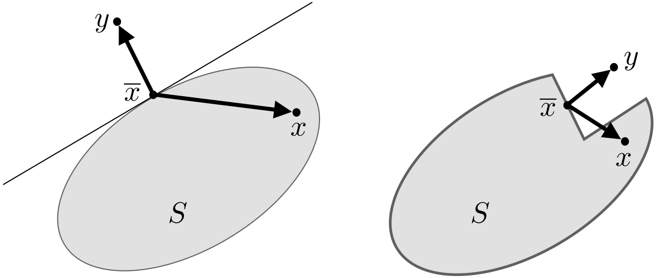
Given a convex set S and a point $y \notin S$, we want to find $\overline{x} \in S$ closest to y and a hyperplane that separates them.

Theorem 7 (Closest-point theorem)

Let $S \neq \emptyset$ be a closed convex set in $\mathbb{R}^n$ and $y \notin S$. Then, there exists a unique point $\overline{x} \in S$ with minimum distance from y. In addition, $\overline{x}$ is the minimising point if and only if

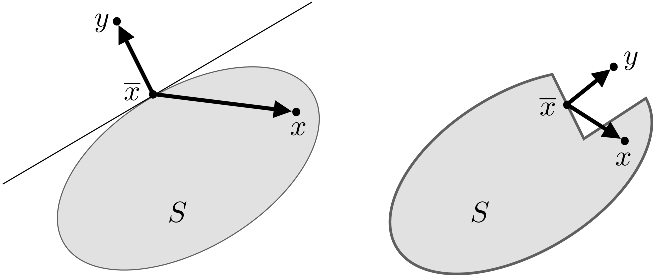
$$(y - \overline{x})^\top (x - \overline{x}) \leq 0, \text{ for all } x \in S$$

Separation and support of sets



Closest-point theorem for a closed convex set (on the left). On the right, an illustration on how the absence of convexity invalidates the result.

Separation and support of sets



Closest-point theorem for a closed convex set (on the left). On the right, an illustration on how the absence of convexity invalidates the result.

Some geometric facts:

- $(y - \overline{x})^\top (x - \overline{x}) \leq 0$ implies that the angle between $(y - \overline{x})$ and $(x - \overline{x})$ is always greater than or equal to 90°.
- S lies in the half-space $p^\top (x - \overline{x}) \leq 0$ relative to the hyperplane $p^\top (x - \overline{x}) = 0$, with normal $p = (y - \overline{x})$.

Separation of two sets and hyperplanes

Definition 8 (Separation of sets)

Let S_1 and S_2 be nonempty sets in $\mathbb{R}^n$. $H = \{x : p^\top x = \alpha\}$ is said to (properly) separate S_1 and S_2 if $p^\top x \geq \alpha$ for each $x \in S_1$ and $p^\top x \leq \alpha$ for each $x \in S_2$ (and $S_1 \cup S_2 \not\subset H$). In addition, we have:

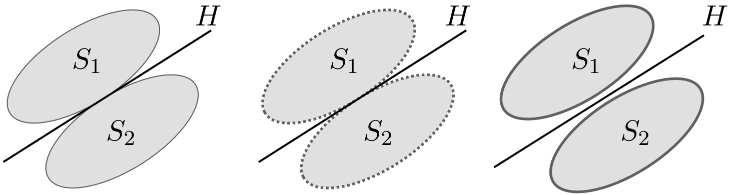
- **Strict separation:** $p^\top x < \alpha$ for each $x \in S_1$ and $p^\top x > \alpha$ for each $x \in S_2$;
- **Strong separation:** $p^\top x \geq \alpha + \epsilon$ for some $\epsilon > 0$ and $x \in S_1$, and $p^\top x \leq \alpha$ for each $x \in S_2$.

Separation of two sets and hyperplanes

Definition 8 (Separation of sets)

Let S_1 and S_2 be nonempty sets in $\mathbb{R}^n$. $H = \{x : p^\top x = \alpha\}$ is said to (properly) separate S_1 and S_2 if $p^\top x \geq \alpha$ for each $x \in S_1$ and $p^\top x \leq \alpha$ for each $x \in S_2$ (and $S_1 \cup S_2 \not\subset H$). In addition, we have:

- **Strict separation:** $p^\top x < \alpha$ for each $x \in S_1$ and $p^\top x > \alpha$ for each $x \in S_2$;
- **Strong separation:** $p^\top x \geq \alpha + \epsilon$ for some $\epsilon > 0$ and $x \in S_1$, and $p^\top x \leq \alpha$ for each $x \in S_2$.



Three types of separation between S_1 and S_2.

Separation of a set and a point

The following separation theorem is a fundamental result from which many other results will follow.

Theorem 9 (Separation theorem)

Let $S \neq \emptyset$ be a closed convex set in $\mathbb{R}^n$ and $y \notin S$. Then, there exists a nonzero vector $p \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that $p^\top x \leq \alpha$ for each $x \in S$ and $p^\top y > \alpha$.

Separation of a set and a point

The following separation theorem is a fundamental result from which many other results will follow.

Theorem 9 (Separation theorem)

Let $S \neq \emptyset$ be a closed convex set in $\mathbb{R}^n$ and $y \notin S$. Then, there exists a nonzero vector $p \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that $p^\top x \leq \alpha$ for each $x \in S$ and $p^\top y > \alpha$.

Proof.

Theorem 7 guarantees the existence of a unique minimising $\overline{x} \in S$ such that $(y - \overline{x})^\top (x - \overline{x}) \leq 0$ for each $x \in S$. Let $p = (y - \overline{x}) \neq 0$ and $\alpha = \overline{x}^\top (y - \overline{x}) = p^\top \overline{x}$. Then we get $p^\top x \leq \alpha$ for each $x \in S$, while $p^\top y - \alpha = (y - \overline{x})^\top (y - \overline{x}) = ||y - \overline{x}||^2 > 0$. □

Separation of a set and a point

Remark: interesting consequences of Theorem 9:

- if S is closed and convex, then it is the intersection of all half-spaces containing S.
- if $y \notin \mathbf{clo}(\mathbf{conv}(S))$, then strong separation holds.

Separation of a set and a point

Remark: interesting consequences of Theorem 9:

- if S is closed and convex, then it is the intersection of all half-spaces containing S.
- if $y \notin \mathbf{clo}(\mathbf{conv}(S))$, then strong separation holds.

Theorem 9 plays a central role in the famous Farkas' Theorem.

- states that a specific system of (in)equalities has a solution if a related second system does not.

Separation of a set and a point

Remark: interesting consequences of Theorem 9:

- if S is closed and convex, then it is the intersection of all half-spaces containing S.
- if $y \notin \mathbf{clo}(\mathbf{conv}(S))$, then strong separation holds.

Theorem 9 plays a central role in the famous Farkas' Theorem.

- states that a specific system of (in)equalities has a solution if a related second system does not.
- Fundamental for deriving optimality conditions and infeasibility certificates in LP problems.

Separation of a set and a point

Remark: interesting consequences of Theorem 9:

- if S is closed and convex, then it is the intersection of all half-spaces containing S.
- if $y \notin \mathbf{clo}(\mathbf{conv}(S))$, then strong separation holds.

Theorem 9 plays a central role in the famous Farkas' Theorem.

- states that a specific system of (in)equalities has a solution if a related second system does not.
- Fundamental for deriving optimality conditions and infeasibility certificates in LP problems.
- Several variants referred to as Farkas' lemma can be found in the optimisation literature.

Farkas' theorem

Theorem 10

Let A be an $m \times n$ matrix and c be an n-vector. Then exactly one of the following two systems has a solution:

$$(1): \; Ax \leq 0, c^\top x > 0, x \in \mathbb{R}^n$$

$$(2): \; A^\top y = c, y \geq 0, y \in \mathbb{R}^m.$$

Farkas' theorem

Theorem 10

Let A be an $m \times n$ matrix and c be an n-vector. Then exactly one of the following two systems has a solution:

$$(1): \; Ax \leq 0, c^\top x > 0, x \in \mathbb{R}^n$$
$$(2): \; A^\top y = c, y \geq 0, y \in \mathbb{R}^m.$$

Proof.

Suppose (2) has a solution. Let x be such that $Ax \leq 0$. Then $c^\top x = (A^\top y)^\top x = y^\top Ax \leq 0$. Hence, (1) has no solution.

Farkas' theorem

Theorem 10

Let A be an $m \times n$ matrix and c be an n-vector. Then exactly one of the following two systems has a solution:

$$(1): \; Ax \leq 0, c^\top x > 0, x \in \mathbb{R}^n$$
$$(2): \; A^\top y = c, y \geq 0, y \in \mathbb{R}^m.$$

Proof.

Suppose (2) has a solution. Let x be such that $Ax \leq 0$. Then $c^\top x = (A^\top y)^\top x = y^\top Ax \leq 0$. Hence, (1) has no solution.

Next, suppose (2) has no solution. Let $S = \{x : x = A^\top y, y \geq 0\}$. Notice that S is closed and convex and that $c \notin S$. By Theorem 9, there exists $p \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that $p^\top c > \alpha$ and $p^\top x \leq \alpha$ for $x \in S$.

As $0 \in S$, $\alpha \geq 0$ and $p^\top c > 0$. Also, $\alpha \geq p^\top A^\top y = y^\top Ap$ for $y \geq 0$. This implies that $Ap \leq 0$, and thus p satisfies (1). □

Geometry of the Farkas' theorem

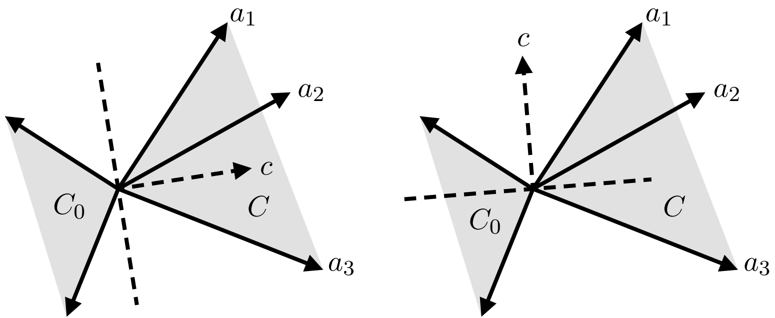
Consider the cone formed by the rows a_i of A:
$C = \left\{c \in \mathbb{R}^n : c_j = \sum_{i=1}^{m} a_{ij} y_i, j = 1, \ldots, n, y_i \geq 0, i = 1, \ldots, m\right\}.$

Its polar cone is given by $C^0 = \{x : Ax \leq 0\}$.

Geometry of the Farkas' theorem

Consider the cone formed by the rows a_i of A:
$C = \{c \in \mathbb{R}^n : c_j = \sum_{i=1}^m a_{ij}y_i, j = 1, \dots, n, y_i \geq 0, i = 1, \dots, m\}$.

Its polar cone is given by $C^0 = \{x : Ax \leq 0\}$. If $c \in C$, then (2) has a solution. Otherwise, (1) has a solution as $\left\{x : c^\top x > 0\right\} \cap C^0 \neq \emptyset$.



Geometrical illustration of the Farkas' theorem. On the left, system (2) has a solution, while on the right, system (1) has a solution

Supporting of sets at boundary points

Definition 11 (Supporting hyperplane)

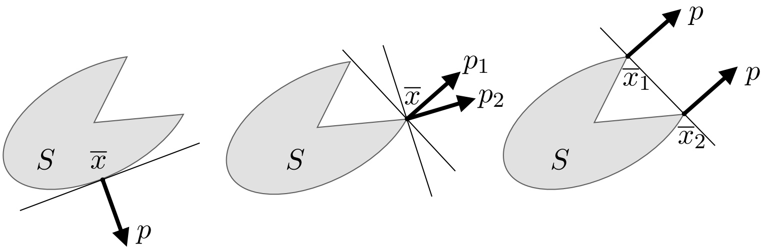
Let $H = \{x \in \mathbb{R}^n : p^\top(x - \overline{x}) = 0\}$, $S \neq \emptyset$ be a set in $\mathbb{R}^n$, and $\overline{x} \in \mathbf{bou}(S)$. H is a supporting hyperplane of S at $\overline{x}$ if either $S \subseteq H^+$ (i.e., $p^\top(x - \overline{x}) \geq 0$ for $x \in S$) or $S \subseteq H^-$.

Supporting of sets at boundary points

Definition 11 (Supporting hyperplane)

Let $H = \{x \in \mathbb{R}^n : p^\top(x - \overline{x}) = 0\}$, $S \neq \emptyset$ be a set in $\mathbb{R}^n$, and $\overline{x} \in \mathbf{bou}(S)$. H is a supporting hyperplane of S at $\overline{x}$ if either $S \subseteq H^+$ (i.e., $p^\top(x - \overline{x}) \geq 0$ for $x \in S$) or $S \subseteq H^-$.

Note that H is a supporting hyperplane if $p^\top\overline{x} = \inf\{p^\top x : x \in S\}$ (or $p^\top\overline{x} = \sup\{p^\top x : x \in S\}$).



Supporting hyperplanes for an arbitrary set.

Supporting of sets at boundary points

An important characteristic of convex sets that we will use is that they have supporting hyperplanes for all boundary points.

Theorem 12 (Support of convex sets)

Let $S \neq \emptyset$ *be a convex set in* $\mathbb{R}^n$*, and let* $\overline{x} \in \mathbf{bou}(S)$*. Then there exists* $p \neq 0^n$ *such that* $p^\top(x - \overline{x}) \leq 0$ *for each* $x \in \mathbf{clo}(S)$*.*

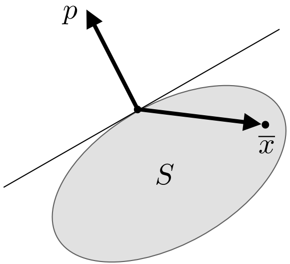
Supporting of sets at boundary points

An important characteristic of convex sets that we will use is that they have supporting hyperplanes for all boundary points.

Theorem 12 (Support of convex sets)

Let $S \neq \emptyset$ be a convex set in $\mathbb{R}^n$, and let $\overline{x} \in \mathbf{bou}(S)$. Then there exists $p \neq 0^n$ such that $p^\top(x - \overline{x}) \leq 0$ for each $x \in \mathbf{clo}(S)$.

The proof follows from Theorem 9, without explicitly considering an $y \notin S$ and by noticing that $\mathbf{bou}(S) \subset \mathbf{clo}(S)$.



Supporting hyperplanes for convex sets. Notice how every boundary point has at least one supporting hyperplane


THANK YOU

Reference

Sahinidis, Nikolaos V. "Mixed-integer nonlinear programming 2018." Optimization and Engineering 20 (2019): 301-306.

Floudas, Christodoulos A. Nonlinear and mixed-integer optimization: fundamentals and applications. Oxford University Press, 1995.

Tawarmalani, Mohit, and Nikolaos V. Sahinidis. Convexification and global optimization in continuous and mixed-integer nonlinear programming: theory, algorithms, software, and applications. Vol. 65. Springer Science & Business Media, 2013.