ORF 522

Linear Programming and Convex Analysis

Basics of Convexity

Today

- A few elementary definitions about convexity,
- Extreme points,
- Separating and supporting hyperplanes,
- Carathéodory Theorem.

Reminder: Basic solutions and hyperplanes

- When $A\mathbf{x} = \mathbf{b}$, $A \in \mathbf{R}^{m\times n}$, $\mathbf{Rank}(A) = m < n$ then,
  - we can **choose a list $\mathbf{I}$** of m **basic variables** among n,
  - solutions such that $\mathbf{x}_i = 0$ for $i \notin \mathbf{I}$ are called basic,
  - When $\mathbf{b}$ is l.i. from any subset of $m-1$ columns of $B_{\mathbf{I}}$ then the $x_i \neq 0, i \in \mathbf{I}$ and the solution is **not degenerate**.

- the set $H_{\mathbf{c},z} = \left\{\mathbf{x} \in \mathbf{R}^n | \mathbf{c}^T\mathbf{x} = z\right\}$, $\mathbf{c} \neq \mathbf{0}$ is a hyperplane
  - $\mathbf{c}$ is a **normal vector** to the hyperplane,
  - The vector subspace $H_{\mathbf{c},0}$ and the affine spaces $H_{\mathbf{c},z}$ are **parallel**.
  - Given a hyperplane H we define open halfspaces H_+ and H_- and their closures $\overline{H_+}$ and $\overline{H_-}$.

"In response to..." Short comment about degeneracy

- Degeneracy only means something for a **linear equation**. **NO inequalities yet**
- Simple example in $\mathbf{R}^3$. We can't draw picture beyond.

$$\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_3 = 1 \end{cases}$$

- $A\mathbf{x} = \mathbf{b}$.

$$A_b = \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{array}\right]$$

- All groups of 2 columns of A are l.i. hence we have three basic solutions. i.e. Solutions where we want to control zero patterns.
- $[\mathbf{0}, ?, ?], [?, \mathbf{0}, ?], [?, ?, \mathbf{0}]$.
- In fact, $[\mathbf{0}, 1, 1], [1, \mathbf{0}, \mathbf{0}], [1, \mathbf{0}, \mathbf{0}]$.
- Two basic solutions with the same value... not very satisfying.

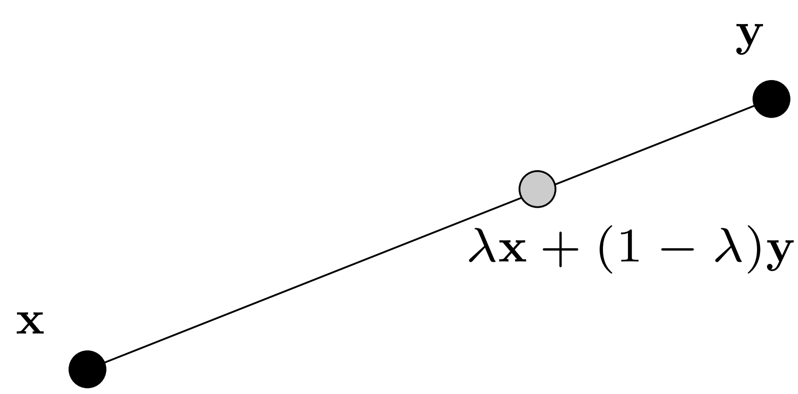
"In response to..." Short comment about degeneracy

- Let's try with inequalities.
- A canonical program with two variables x_1, x_2. We start in $\mathbf{R}^2$
- Add 4 inequalities (assume $\geq$) add 4 slack variables. The problem is in $\mathbf{R}^6$.
- We have 4 vectors of $\mathbf{R}^6$, the **rows** of A.
- A **non-degenerate** basic solution has **4 non-zero components**. **2 are zero.**
- set variables 1 and 2 at zero. unless a hyperplane cuts the origin, no degeneracy
- set one of variables 1 or 2 at zero. The other must be crossing a hyperplane.
- set 1 & 2 be non zero. Let's look for degeneracy.
- we find out that this means 3 lines have a common point
- actually that means the two first columns and the b column are tied, l.d.

Convex sets & extreme points

Definition

- Convexity starts by defining segments



$$[\mathbf{x}, \mathbf{y}] = \lambda\mathbf{x} + (1-\lambda)\mathbf{y}, \lambda \in [0, 1]$$

.

Definition 1. *A set C is said to be* **convex** *if for all $\mathbf{x}$ and $\mathbf{y}$ in C the segment $[\mathbf{x}, \mathbf{y}] \subset C$.*

Examples

- $\mathbf{R}^n$ is trivially convex and so is any vector subspace V of $\mathbf{R}^n$.
- For $\mathbf{R}^n \ni \mathbf{c} \neq \mathbf{0}$ and $z \in \mathbf{R}$, $H_{\mathbf{c},z}$ is convex
- The halfspaces $H^+_{\mathbf{c},z}$ and $H^-_{\mathbf{c},z}$ are **open convex sets**, their respective closures are **closed convex sets**.
- Let $\mathbf{x}_1, \mathbf{x}_2 \in B_r(\mathbf{x}_0), \lambda \in [0,1]$ then

$$|(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) - \mathbf{x}_0| = |\lambda(\mathbf{x}_1 - \mathbf{x}_0) + (1-\lambda)(\mathbf{x}_2 - \mathbf{x}_0)| < \lambda r + (1-\lambda) r = r.$$

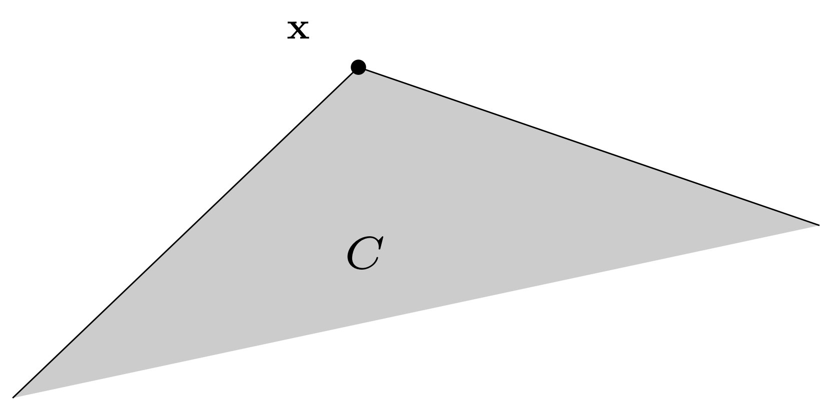
hence $B_r(\mathbf{x}_0)$ and similarly $\overline{B_r(\mathbf{x}_0)}$ are convex

Extreme points

Definition 2. *A point* $\mathbf{x}$ *of a convex set* C *is said to be an* ***extreme point*** *of* C *if*

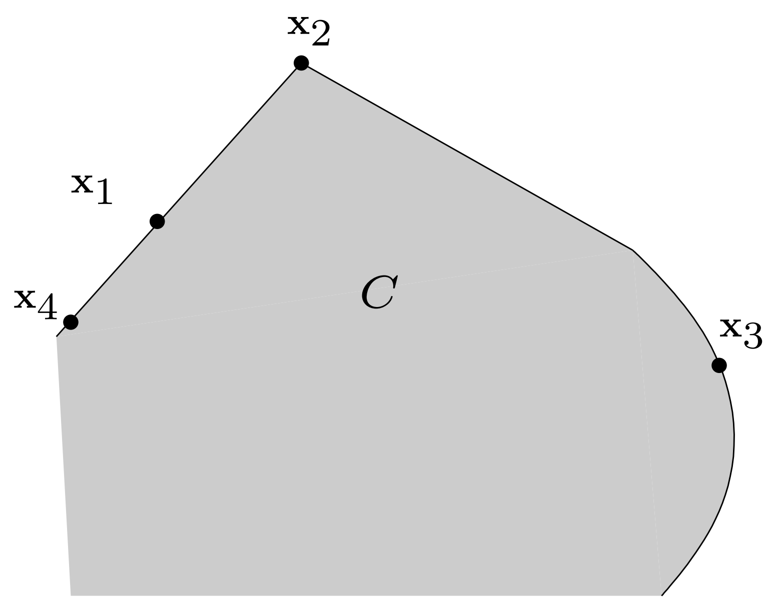
$$\left(\exists \mathbf{x}_1, \mathbf{x}_2 \in C \mid \mathbf{x} = \tfrac{\mathbf{x}_1 + \mathbf{x}_2}{2}\right) \Rightarrow \mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}.$$

- intuitively $\mathbf{x}$ is not part of an **open** segment of two other points $\mathbf{x}_1, \mathbf{x}_2$.
- other definitions use $0 < \lambda < 1, \mathbf{x} = \lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2$ but the one above is equivalent & easier to remember.



Extreme points

- an extreme point is a boundary point but **the converse is not true in general**.



- $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ are all boundary points. Only $\mathbf{x}_2$ and $\mathbf{x}_3$ are extreme. $\mathbf{x}_1$ for instance can be written as $\lambda\mathbf{x}_2 + (1-\lambda)\mathbf{x}_4$

Hyperplanes and Convexity: Isolation and Support

Boundaries of Hyperplanes and Halfspaces

- Hyperplanes are closed
  - We can actually show that $H_{\mathbf{c},z} \subset \partial H_{\mathbf{c},z}$, namely any point of $H_{\mathbf{c},z}$ is a boundary point:
    - let $\mathbf{x} \in H_{\mathbf{c},z}$ and $B_r(\mathbf{x})$ an open ball centered in $\mathbf{x}$.
    - let $\mathbf{y}_1 = \mathbf{x} + \frac{r}{2|\mathbf{c}|^2}\mathbf{c}$. Then $\mathbf{c}^T\mathbf{y}_1 = z + \frac{r}{2} > z$ hence $\mathbf{y}_1 \notin H_{\mathbf{c},z}$ but $\mathbf{y}_1 \in B_r(\mathbf{x})$,
    - let $\mathbf{z} \in H_{\mathbf{c},z}, \mathbf{z} \neq \mathbf{x}$, and $\mathbf{y}_2 = \mathbf{x} + r\frac{\mathbf{x}-\mathbf{z}}{2|\mathbf{x}-\mathbf{z}|}$, hence $\mathbf{y}_2 \in H_{\mathbf{c},z}$ and $\mathbf{y}_2 \in B_r(\mathbf{x})$.
  - We could also have raised the fact that for $\mathbf{x}_i$ a converging sequence of $H_{\mathbf{c},z}$ we have that $\mathbf{c}^T \lim_{i\to\infty} \mathbf{x}_i = \lim_{i\to\infty} \mathbf{c}^T\mathbf{x}_i = z$.
- The boundary of a halfspace is the corresponding hyperplane, i.e.

$$\partial H_- = \partial H_+ = H.$$

- The interior H^o of a hyperplane is empty as $H^o = H \setminus \partial H$.

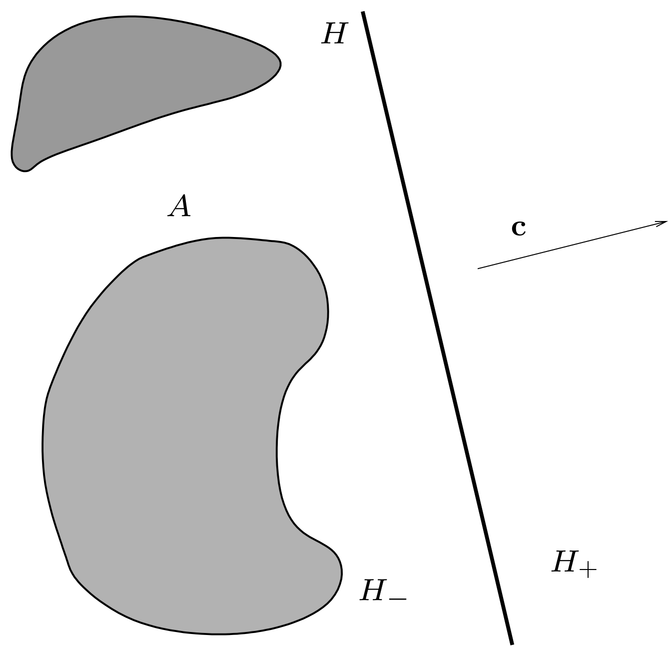
Hyperplanes, halfspaces and convexity

Lemma 1. *(i) All hyperplanes are convex;*
(ii) The halfspaces $H^+_{\mathbf{c},z}, H^-_{\mathbf{c},z}, \overline{H^+_{\mathbf{c},z}}, \overline{H^-_{\mathbf{c},z}}$ *are convex;*
(iii) Any ***intersection*** *of convex sets is convex;*
(iv) The set of all ***feasible solutions of a linear program*** *is a* ***convex*** *set.*

Proof. (i) $c^T(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) = (\lambda + (1-\lambda))\, z = z$.
(ii) same as above by replacing equality by inequalities.
(iii) Let $C = \cap_{i\in I} C_i$. Let $\mathbf{x}_1, \mathbf{x}_2 \in C$. Then for $\lambda \in [0,1], \forall i \in I, (\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \in C_i$, hence $(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \in C$.
(iv) The set of feasible points to an LP problem is the intersection of hyperplanes $\mathbf{r}_i^T\mathbf{x} = b_i$ and halfspaces $\mathbf{r}_j^T\mathbf{x} \gtreqless b_j$ and is hence convex by (iii). ■

Isolation

Definition 3. *Let $A \subset \mathbf{R}^n$ be a set and let $H \subset \mathbf{R}^n$ be a affine hyperplane. H is said to* ***isolate*** *A if A is contained in one of the closed subspaces $\overline{H_-}$ or $\overline{H_+}$. H* ***strictly isolates*** *A if A is contained in one of the open halfspaces H_- or H_+.*



Isolation Theorem

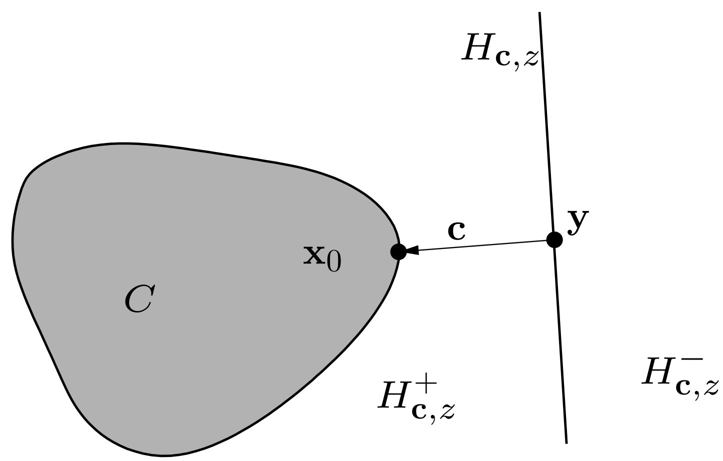
Theorem 1. *Let C be a **closed convex set** and $\mathbf{y}$ a point not in C. Then there is a hyperplane $H_{\mathbf{c},z}$ that contains $\mathbf{y}$ and such that $C \subset H^{-}_{\mathbf{c},z}$ or $C \subset H^{+}_{\mathbf{c},z}$*

- (Bar02,II.1.6) has a more general result when C is **open**. The proof is longer and we won't use it.
- **Proof strategy**: build a suitable hyperplane and show it satisfies the property.

Isolation Theorem : Proof

Proof. • **Define the hyperplane**:

- Let $\delta = \inf_{x \in C} |\mathbf{x} - \mathbf{y}| > 0$.
- The continuous function $\mathbf{x} \to |\mathbf{x} - \mathbf{y}|$ on the closed set $\overline{B_{2\delta}(\mathbf{y})}$ achieves its minimum at a point $\mathbf{x}_0 \in C$.
- One can prove that necessarily $\mathbf{x} \in \partial C$.
- Let $\mathbf{c} = \mathbf{x}_0 - \mathbf{y}, z = \mathbf{c}^T\mathbf{y}$ and consider $H_{\mathbf{c},z}$. Clearly $\mathbf{y} \in H_{\mathbf{c},z}$.



Isolation Theorem : Proof

- **Show that** $C \subset H^{+}_{\mathbf{c},z}$:
    - Let $\mathbf{x} \in C$. Since $\mathbf{x}_0 \in C$, for $\lambda \in [0,1]$,

$$\lambda\mathbf{x} + (1-\lambda)\mathbf{x}_0 = \mathbf{x}_0 + \lambda(\mathbf{x}-\mathbf{x}_0) \in C.$$

    - By definition of $\mathbf{x}_0$, $|\,(\mathbf{x}_0 + \lambda(\mathbf{x}-\mathbf{x}_0)) - \mathbf{y}|^2 \geq |\mathbf{x}_0 - \mathbf{y}|^2$,
    - thus by definition of $\mathbf{c} = \mathbf{x}_0 - \mathbf{y}$,

$$|\lambda(\mathbf{x}-\mathbf{x}_0) + \mathbf{c}|^2 \geq |\mathbf{c}|^2,$$

    - thus $2\lambda\mathbf{c}^T(\mathbf{x}-\mathbf{x}_0) + \lambda^2|\mathbf{x}-\mathbf{x}_0|^2 \geq 0$,
    - Letting $\lambda \to 0$ we have that $\mathbf{c}^T(\mathbf{x}-\mathbf{x}_0) \geq 0$, hence

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{c}^T\mathbf{x}_0 = \mathbf{c}^T(\mathbf{y}+\mathbf{c}) = z + |\mathbf{c}|^2 = z + \delta^2 > z$$

■

Supporting Hyperplane

Definition 4. *Let* $\mathbf{y}$ *be a* ***boundary*** *point of a convex set* C*. A hyperplane* $H_{\mathbf{c},z}$ *is called a* ***supporting hyperplane*** *of* C *at* $\mathbf{y}$ *if* $\mathbf{y} \in H_{\mathbf{c},z}$ *and either* $C \subseteq \overline{H^{+}_{\mathbf{c},z}}$ *or* $C \subseteq \overline{H^{-}_{\mathbf{c},z}}$*.*

Theorem 2. *If* $\mathbf{y}$ *is a boundary point of a closed convex set* C *then there is at least one supporting hyperplane at* $\mathbf{y}$*.*

- **Proof strategy**: use the isolation theorem on a sequence of points that converge to a boundary point.

Supporting Hyperplane : Proof

Proof. Since $\mathbf{y} \in \partial C, \forall k \in \mathbf{N}$, $\exists \mathbf{y}_k \in B_{\frac{1}{k}}(\mathbf{y})$ such that $\mathbf{y}_k \notin C$. $(\mathbf{y}_k)$ is thus a sequence of $\mathbf{R}^n \setminus C$ that converges to $\mathbf{y}$. Let $\mathbf{c}_k$ be the sequence of corresponding normal vectors constructed according to the proof of Theorem 1, normalized so that $|\mathbf{c}_k| = 1$ and C is in the halfspace $\{\mathbf{x} \,|\, \mathbf{c}_k^T \mathbf{x} \geq \mathbf{c}_k^T \mathbf{y}_k\}$. Since $(\mathbf{c}_k)$ is a bounded sequence in a compact space, there exists a subsequence $\mathbf{c}_{k_j}$ that converges to a point $\mathbf{c}$. Let $z = \mathbf{c}^T \mathbf{y}$. For any $\mathbf{x} \in C$,

$$\mathbf{c}^T \mathbf{x} = \lim_{j \to \infty} \mathbf{c}_{k_j}^T \mathbf{x} \geq \lim_{j \to \infty} \mathbf{c}_{k_j}^T \mathbf{y}_{k_j} = \mathbf{c}^T \mathbf{y} = z,$$

thus $C \subset \overline{H_{\mathbf{c},z}^{+}}$ ■

Bounded from below

Definition 5. *A set* $A \subset \mathbf{R}^n$ *is said to be* ***bounded from below*** *if for all* $1 \leq j \leq n$,

$$\inf \left\{ \mathbf{x}_j \,|A \ni \mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T \right\} > -\infty.$$

- Any bounded set is bounded from below
- More importantly, $\mathbf{R}^n_+ = \{\mathbf{x} | \mathbf{x} \geq 0\}$ is bounded from below.
- the LP set of solutions $\{\mathbf{x} \in \mathbf{R}^n \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ is **convex** & **bounded from below**.

Supporting Hyperplane and Extreme Points

Theorem 3. *Let C be a closed convex set which is bounded from below. Then every supporting hyperplane of C contains an extreme point of C .*

- **Proof strategy**: Show that for a supporting hyperplane H, an extreme point of the convex subset $H \cap C$ is an extreme point of C. Find an extreme point of $H \cap C$.

Supporting Hyperplane and Extreme Points: Proof

Proof. • Let $H_{\mathbf{c},z}$ be a supporting hyperplane at $\mathbf{y} \in C$. Let us write $A = H_{\mathbf{c},z} \cap C$ which is non-empty since it contains $\mathbf{y}$.

• **an extreme point of A is an extreme point of C**

- suppose $\mathbf{x} \in A$, that is $\mathbf{c}^T\mathbf{x} = z$, is **not** an ext. point of C, i.e $\exists \mathbf{x}_1 \neq \mathbf{x}_2 \in C$ such that $\mathbf{x} = \frac{\mathbf{x}_1+\mathbf{x}_2}{2}$.
- If $\mathbf{x}_1 \notin A$ **or** $\mathbf{x}_2 \notin A$ then $\frac{1}{2}\mathbf{c}^T(\mathbf{x}_1 + \mathbf{x}_2) > z = \mathbf{c}^T\mathbf{x}$ hence $\mathbf{x}_1, \mathbf{x}_2 \in A$ and thus $\mathbf{x}$ is **not** an ext. point of A.

Supporting Hyperplane and Extreme Points: Proof

- **look now for an extreme point of A. We use mainly $A \subset H_{\mathbf{c},z} \cap \mathbf{R}_+^m$**
  - if A is a singleton, namely $A = \{\mathbf{y}\}$, then $\mathbf{y}$ is obviously extreme.
  - if not, **narrow down recursively**:
    - $A^1 = \operatorname{argmin}\{\mathbf{a}_1 \mid \mathbf{a} \in A\}$. Since $A \subset C$ and C is bounded from below the closed set A^1 is well defined as the set of points which achieve this minimum.
    - If A^1 is still not a singleton, we narrow further:

$$A^j = \operatorname{argmin}\{\mathbf{a}_j \mid \mathbf{a} \in A^{j-1}\}.$$

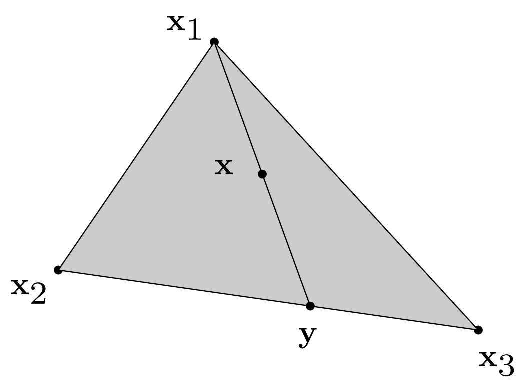
    - Since $A \subset \mathbf{R}^n$, this process must stop after $k \leq n$ iterations (after n iterations the n variables of points in A^n are uniquely defined). We have $A^k \subseteq A^{k-1} \subseteq A^1 \subseteq A$ and write $A^k = \{\mathbf{a}^k\}$.
  - Suppose $\exists \mathbf{x}^1 \neq \mathbf{x}^2 \in A$ such that $\mathbf{a}^k = \frac{\mathbf{x}^1+\mathbf{x}^2}{2}$. In particular $\forall i \leq k, \mathbf{a}_i^k = \frac{\mathbf{x}_i^1+\mathbf{x}_i^2}{2}$.
  - Since $\mathbf{a}_1^k$ is an **infimum**, $\mathbf{x}_i^1 = \mathbf{x}_i^2 = \mathbf{a}_1^k$ and $\mathbf{x}^1, \mathbf{x}^2 \in A^1$.
  - By the same argument **applied recursively** we have that $\mathbf{x}^1, \mathbf{x}^2 \in A^j$ and finally A^k which by construction is $\{\mathbf{a}^k\}$, hence $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{a}^k$, a contradiction, and $\mathbf{a}^k$ is our extreme point.

■

Convex Hulls & Carathéodory's Theorem

Convex combinations

Definition 6. *Let* $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\}$ *be a set of points. Let* $\alpha_1, \cdots, \alpha_k$ *be a family of nonnegative weights such that* $\sum_1^k \alpha_i = 1$*. Then* $\mathbf{x} = \sum_1^k \alpha_i \mathbf{x}_i$ *is called a* ***convex combination*** *of the points* $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$*.*



Let's illustrate this statement with a point $\mathbf{x}$ in a triangle $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$.

- Let $\mathbf{y}$ be the intersection of $(\mathbf{x}_1, \mathbf{x})$ with $[\mathbf{x}_2, \mathbf{x}_3]$. $\mathbf{y} = p\mathbf{x}_2 + q\mathbf{x}_3$ with $p = \frac{|\mathbf{x}_2 - \mathbf{y}|}{|\mathbf{x}_3 - \mathbf{x}_2|}$ and $q = \frac{|\mathbf{x}_3 - \mathbf{y}|}{|\mathbf{x}_3 - \mathbf{x}_2|}$.
- On the other hand, $\mathbf{x} = l\mathbf{x}_1 + k\mathbf{y}$ with $l = \frac{|\mathbf{x}_1 - \mathbf{x}|}{|\mathbf{x}_1 - \mathbf{y}|}$ and $k = \frac{|\mathbf{y} - \mathbf{x}|}{|\mathbf{x}_1 - \mathbf{y}|}$.
- Finally $\mathbf{x} = l\mathbf{x}_1 + pk\mathbf{x}_2 + qk\mathbf{x}_3$, and $l + pk + qk = 1$.

Convex hull

Definition 7. *The* **convex hull** $\langle A \rangle$ *of a set* A *is the minimal convex set that contains* A.

Lemma 2. *(i) if* $A \neq \emptyset$ *then* $\langle A \rangle \neq \emptyset$
(ii) if $A \subset B$ *then* $\langle A \rangle \subset \langle B \rangle$
(iii) $\langle A \rangle$ *is the intersection of all convex sets that contain* A.
(iv) if A *is convex then* $\langle A \rangle = A$

Convex hull $\Leftrightarrow$ all convex combinations

Theorem 4. *The convex hull of a set of points* $\{\mathbf{x}_1, \cdots, \mathbf{x}_k\}$ *is the set of all convex combinations of* $\mathbf{x}_1, \cdots, \mathbf{x}_k$.

Proof. • Let $A = \{\mathbf{x} \,|\, \mathbf{x} = \sum_1^k \alpha_i \mathbf{x}_i, \alpha_i \geq 0, \sum_1^k \alpha_i = 1\}$; $B = \langle\{\mathbf{x}_1, \cdots, \mathbf{x}_k\}\rangle$

- It's easy to prove that A is convex: Let $\mathbf{x} = \sum_1^k \alpha_i \mathbf{x}_i$ and $\mathbf{y} = \sum_1^k \beta_i \mathbf{x}_i$ be two points of A. Then $\lambda\mathbf{x} + (1-\lambda)\mathbf{y}$ can be written as

$$\sum_{i=1}^{k} (\lambda\alpha_i + (1-\lambda)\beta_i)\, \mathbf{x}_i \in A$$

- $B \subseteq A$: A is convex and contains each point $\mathbf{x}_i$ since

$$\mathbf{x}_i = \sum_{j=1}^{k} \delta_{ij}\mathbf{x}_j.$$

Convex hull $\Leftrightarrow$ all convex combinations

- $A \subseteq B$: by induction on k. if $k = 1$ then $B_1 = \langle\{\mathbf{x}_1\}\rangle$ and $A_1 = \{\mathbf{x}_1\}$. By Lemma 2 $A_1 \subseteq B_1$. Suppose that the claim holds for any family of $k-1$ points, i.e. $A_{k-1} \subseteq B_{k-1}$. Let now $\mathbf{x} \in A_k$ such that

$$\mathbf{x} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i.$$

If $\mathbf{x} = \mathbf{x}_k$ then trivially $\mathbf{x} \in B_k$. If $\mathbf{x} \neq \mathbf{x}_k$ then $\alpha_k \neq 1$ and we have that

$$\frac{\sum_{i=1}^{k-1} \alpha_i}{1 - \alpha_k} = 1.$$

Consider $\mathbf{y} = \sum_{i=1}^{k-1} \frac{\alpha_i}{1-\alpha_k} \mathbf{x}_i$. $\mathbf{y} \in B_{k-1}$ by the induction hypothesis. Since $\{\mathbf{x}_1, \cdots, \mathbf{x}_{k-1}\} \subset \{\mathbf{x}_1, \cdots, \mathbf{x}_k\}$, $B_{k-1} \subseteq B_k$ by Lemma2. Since B_k is convex and both $\mathbf{y}, \mathbf{x}_k \in B_k$, so is $\mathbf{x} = (1-\alpha_k)\mathbf{y} + \alpha_k \mathbf{x}_k$.

■

Polytope, Polyhedrons

Definition 8. *The convex hull of a finite set of points in* $\mathbf{R}^n$ *is called a* ***polytope***.

Let $\mathbf{r}_1, \cdots, \mathbf{r}_m$ *be vectors from* $\mathbf{R}^n$ *and* $b_1, \cdots, b_m$ *be numbers. The set*

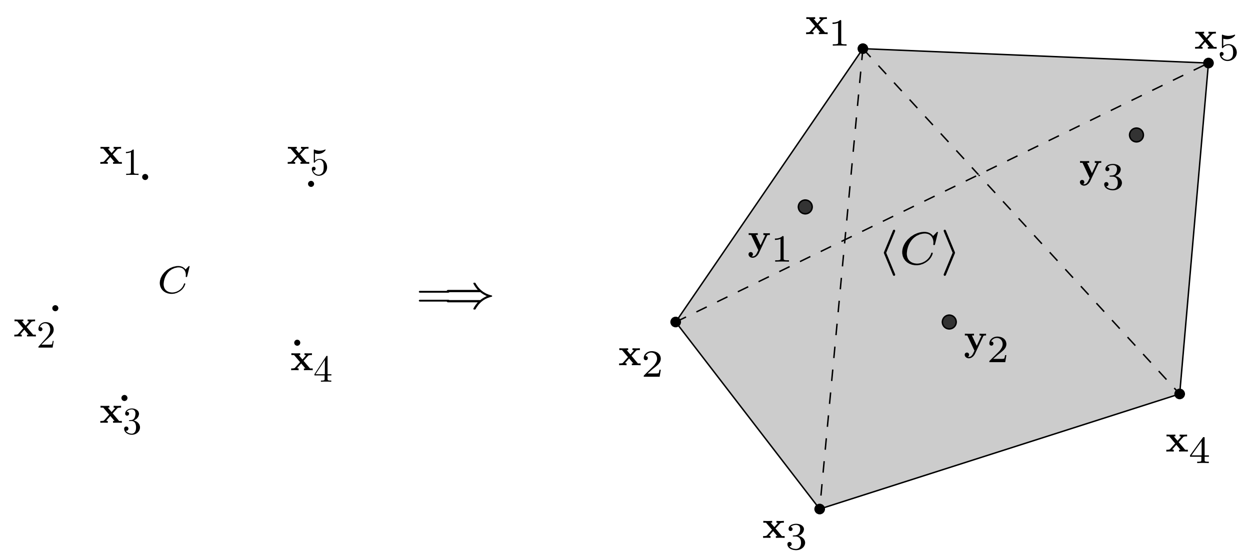
$$P = \left\{ \mathbf{x} \in \mathbf{R}^n \mid \mathbf{r}_i^T \mathbf{x} \leq b_i\ ,\ i = 1, \cdots, n \right\}$$

is called a ***polyhedron***.

- A few comments:
  - **bounded polyhedron** $\Leftrightarrow$ **polytope**: TBP **Weyl-Minkowski** theorem.
  - polytopes are generated by a finite set of points. $\overline{B_r(\mathbf{x})}$ is **not** a polytope.
  - a polyhedron is exactly the set of **feasible solutions of an LP**.

Carathéodory's Theorem

- Start with the example of $C = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5\} \subset \mathbf{R}^2$ and its hull $\langle C \rangle$.



  - $\mathbf{y}_1$ can be written as a convex combination of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ (or $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5$);
  - $\mathbf{y}_2$ can be written as a convex combination of $\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4$;
  - $\mathbf{y}_3$ can be written as a convex combination of $\mathbf{x}_1, \mathbf{x}_4, \mathbf{x}_5$;

- For a set C of 5 points in $\mathbf{R}^{\mathbf{2}}$ there seems to be always a way to write a point $\mathbf{y} \in \langle C \rangle$ as the convex combination of $\mathbf{2+1=3}$ of such points.

- Is this result still valid for **general hulls** $\langle S \rangle$ (not necessarily polytopes but also balls etc..) and **higher dimensions**?

Carathéodory's Theorem

Theorem 5. *Let $S \subset \mathbf{R}^{n}$. Then every point $\mathbf{x}$ of $\langle S \rangle$ can be represented as a convex combination of $n+1$ points from S,*

$$\mathbf{x} = \alpha_1 \mathbf{x}_1 + \cdots + \alpha_{n+1} \mathbf{x}_{n+1}, \sum_{i=1}^{n+1} \alpha_i = 1, \alpha_i \geq 0.$$

alternative formulation:

$$\langle S \rangle = \bigcup_{C \subset S, \mathrm{card}(C) = n+1} \langle C \rangle.$$

- **Proof strategy**: show that when a point is written as a combination of m points and $m > n+1$, it is possible to write it as a combination of $m-1$ points by solving a homogeneous linear equation of $n+1$ equations in $\mathbf{R}^{m}$.

Proof.

- ($\supset$) is direct.
- ($\subset$) any $\mathbf{x} \in \langle S \rangle$ can be written as a convex combination of p points, $\mathbf{x} = \alpha_1 \mathbf{x}_1 + \cdots \alpha_p \mathbf{x}_p$. We can assume $\alpha_i > 0$ for $i = 1, \cdots, p$.
  - If $p < n+1$ then we add terms $0\mathbf{x}_{p+1} + 0\mathbf{x}_{p+2} + \cdots$ to get $n+1$ terms.
  - If $p > n+1$, we build a new combination with one term less:
    - ▷ let $A = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_m \\ 1 & 1 & \cdots & 1 \end{bmatrix} \in \mathbf{R}^{n+1 \times p}$.
    - ▷ **The key here is that since $p > n+1$ there exists a solution $\eta \in \mathbf{R}^m \neq \mathbf{0}$ to $A\eta = \mathbf{0}$.**
    - ▷ By the last row of A, $\eta_1 + \eta_2 + \cdots + \eta_m = 0$, thus η has both $+$ and - coordinates.
    - ▷ Let $\tau = \min\{\frac{\alpha_i}{\eta_i}, \eta_i > 0\} = \frac{\alpha_{i_0}}{\eta_{i_0}}$.
    - ▷ **Let $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{i}} = \boldsymbol{\alpha}_{\boldsymbol{i}} - \boldsymbol{\tau}\boldsymbol{\eta}_{\boldsymbol{i}}$.** Hence $\tilde{\alpha}_i \geq 0$ and $\tilde{\alpha}_{i_0} = 0$.
    - ▷ $$\tilde{\alpha}_1 + \cdots + \tilde{\alpha}_p = (\alpha_1 + \cdots + \alpha_p) - \tau(\eta_1 + \cdots + \eta_p) = 1,$$ $$\tilde{\alpha}_1 \mathbf{x}_1 + \cdots + \tilde{\alpha}_p \mathbf{x}_p = \alpha_1 \mathbf{x}_1 + \cdots + \alpha_p \mathbf{x}_p - \tau(\eta_1 \mathbf{x}_1 + \cdots + \eta_p \mathbf{x}_p) = \mathbf{x}.$$
    - ▷ Thus $\mathbf{x} = \sum_{i \neq i_0} \alpha_i \mathbf{x}_i$ of $\boldsymbol{p - 1}$ points $\{\mathbf{x}_i, i \neq i_0\}$.
    - ▷ Iterate this procedure until $\mathbf{x}$ is a convex combin. of $n+1$ points of S.

Next time

- Some notable points: basic feasible / extreme points / optima
- The simplex in theory