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

• the set $H_{c,z} = \{ x \in \mathbb{R}^n | c^T x = z \}$, $c \neq 0$ is a hyperplane
  
  o $c$ is a **normal vector** to the hyperplane,
  o The vector subspace $H_{c,0}$ and the affine spaces $H_{c,z}$ are **parallel**.
  o 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 \( \mathbb{R}^3 \). We can’t draw picture beyond.

\[
\begin{align*}
&\begin{cases}
  x_1 + x_2 & = 1 \\
x_1 + x_3 & = 1
\end{cases}
\end{align*}
\]

• \( Ax = b \).

\[
A_b = \begin{bmatrix} 1 & 1 & 0 & | & 1 \\ 1 & 0 & 1 & | & 1 \end{bmatrix}
\]

• 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.

• \([0,?,?],[?,0,?],[?,?,0]\).

• In fact, \([0,1,1],[1,0,0],[1,0,0]\).

• Two basic solutions with the same value... not very satisfying.
Let’s try with inequalities.

A canonical program with two variables $x_1, x_2$. We start in $\mathbb{R}^2$.

Add 4 inequalities (assume $\geq$) and add 4 slack variables. The problem is in $\mathbb{R}^6$.

We have 4 vectors of $\mathbb{R}^6$, 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
Convexity starts by defining segments

\[ [x, y] = \lambda x + (1 - \lambda)y, \lambda \in [0, 1] \]

**Definition 1.** A set \( C \) is said to be **convex** if for all \( x \) and \( y \) in \( C \) the segment \( [x, y] \subset C \).
Examples

- \( \mathbb{R}^n \) is trivially convex and so is any vector subspace \( V \) of \( \mathbb{R}^n \).
- For \( \mathbb{R}^n \ni c \neq 0 \) and \( z \in \mathbb{R} \), \( H_{c,z} \) is convex.
- The halfspaces \( H_{c,z}^+ \) and \( H_{c,z}^- \) are open convex sets, their respective closures are closed convex sets.
- Let \( x_1, x_2 \in B_r(x_0), \lambda \in [0,1] \) then

\[
|\lambda x_1 + (1-\lambda)x_2 - x_0| = |\lambda(x_1 - x_0) + (1-\lambda)(x_2 - x_0)| < \lambda r + (1-\lambda) r = r.
\]

hence \( B_r(x_0) \) and similarly \( B_r(x_0) \) are convex.
Definition 2. A point \( x \) of a convex set \( C \) is said to be an extreme point of \( C \) if

\[
(\exists x_1, x_2 \in C \mid x = \frac{x_1 + x_2}{2}) \Rightarrow x_1 = x_2 = x.
\]

- intuitively \( x \) is not part of an open segment of two other points \( x_1, x_2 \).
- other definitions use \( 0 < \lambda < 1, x = \lambda x_1 + (1 - \lambda)x_2 \) 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.**

- $x_1, x_2, x_3, x_4$ are all boundary points. Only $x_2$ and $x_3$ are extreme. $x_1$ for instance can be written as $\lambda x_2 + (1 - \lambda)x_4$
Hyperplanes and Convexity: Isolation and Support
Boundaries of Hyperplanes and Halfspaces

• Hyperplanes are closed
  ○ We can actually show that $H_{c,z} \subset \partial H_{c,z}$, namely any point of $H_{c,z}$ is a boundary point:
    ▶ let $x \in H_{c,z}$ and $B_r(x)$ an open ball centered in $x$.
    ▶ let $y_1 = x + \frac{r}{2|c|^2} c$. Then $c^T y_1 = z + \frac{r}{2} > z$ hence $y_1 \notin H_{c,z}$ but $y_1 \in B_r(x)$.
    ▶ let $z \in H_{c,z}$, $z \neq x$, and $y_2 = x + r \frac{x-z}{2|x-z|}$, hence $y_2 \in H_{c,z}$ and $y_2 \in B_r(x)$.
  ○ We could also have raised the fact that for $x_i$ a converging sequence of $H_{c,z}$ we have that $c^T \lim_{i \to \infty} x_i = \lim_{i \to \infty} c^T x_i = z$.

• 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 $H^o = H \setminus \partial H$. 

ORF-522 12
Lemma 1. (i) All hyperplanes are convex;
(ii) The halfspaces $H_{c,z}^+, H_{c,z}^-$ 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) $c^T(\lambda x_1 + (1 - \lambda)x_2) = (\lambda + (1 - \lambda))z = z$.
(ii) same as above by replacing equality by inequalities.
(iii) Let $C = \cap_{i \in I} C_i$. Let $x_1, x_2 \in C$. Then for
$\lambda \in [0, 1], \forall i \in I, (\lambda x_1 + (1 - \lambda)x_2) \in C_i$, hence $(\lambda x_1 + (1 - \lambda)x_2) \in C$.
(iv) The set of feasible points to an LP problem is the intersection of hyperplanes
$r_i^T x = b_i$ and halfspaces $r_j^T x \geq b_j$ and is hence convex by (iii).
**Definition 3.** Let $A \subset \mathbb{R}^n$ be a set and let $H \subset \mathbb{R}^n$ be an 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 $y$ a point not in $C$. Then there is a hyperplane $H_{c,z}$ that contains $y$ and such that $C \subset H_{c,z}^-$ or $C \subset H_{c,z}^+$. 

- (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.
Proof. • Define the hyperplane:

- Let $\delta = \inf_{x \in C} |x - y| > 0$.
- The continuous function $x \to |x - y|$ on the closed set $B_{2\delta}(y)$ achieves its minimum at a point $x_0 \in C$.
- One can prove that necessarily $x \in \partial C$.
- Let $c = x_0 - y$, $z = c^T y$ and consider $H_{c,z}$. Clearly $y \in H_{c,z}$. 

---

[Diagram of a hyperplane $H_{c,z}$ intersecting a set $C$ with points $x_0$, $c$, and $y$.]
Show that $C \subset H_{c,z}^+$:

- Let $x \in C$. Since $x_0 \in C$, for $\lambda \in [0, 1]$, 
  
  $$
  \lambda x + (1 - \lambda)x_0 = x_0 + \lambda(x - x_0) \in C.
  $$

- By definition of $x_0$, $| (x_0 + \lambda(x - x_0)) - y |^2 \geq |x_0 - y|^2$,
- thus by definition of $c = x_0 - y$,
  
  $$
  |\lambda(x - x_0) + c|^2 \geq |c|^2,
  $$

- thus $2\lambda c^T(x - x_0) + \lambda^2 |x - x_0|^2 \geq 0$,
- Letting $\lambda \to 0$ we have that $c^T(x - x_0) \geq 0$, hence 
  
  $$
  c^T x \geq c^T x_0 = c^T(y + c) = z + |c|^2 = z + \delta^2 > z
  $$

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

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

- **Proof strategy**: use the isolation theorem on a sequence of points that converge to a boundary point.
Proof. Since \( y \in \partial C \), \( \forall k \in \mathbb{N} \), \( \exists y_k \in B_1(y) \) such that \( y_k \notin C \). \((y_k)\) is thus a sequence of \( \mathbb{R}^n \setminus C \) that converges to \( y \). Let \( c_k \) be the sequence of corresponding normal vectors constructed according to the proof of Theorem 1, normalized so that \( |c_k| = 1 \) and \( C \) is in the halfspace \( \{ x \mid c_k^T x \geq c_k^T y_k \} \). Since \((c_k)\) is a bounded sequence in a compact space, there exists a subsequence \( c_{k_j} \) that converges to a point \( c \). Let \( z = c^T y \). For any \( x \in C \),

\[
    c^T x = \lim_{j \to \infty} c_{k_j}^T x \geq \lim_{j \to \infty} c_{k_j}^T y_{k_j} = c^T y = z,
\]

thus \( C \subset \overline{H_{c,z}} \)
Definition 5. A set $A \subset \mathbb{R}^n$ is said to be **bounded from below** if for all $1 \leq j \leq n$,

$$
\inf \{x_j \mid A \ni x = (x_1, \ldots, x_n)^T\} > -\infty.
$$

- Any bounded set is bounded from below
- More importantly, $\mathbb{R}_+^n = \{x \mid x \geq 0\}$ is bounded from below.
- the LP set of solutions $\{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ is convex & bounded from below.
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 $H \cap C$. 


**Proof.**  

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

• **an extreme point of $A$ is an extreme point of $C$**
  
  - suppose $x \in A$, that is $c^T x = z$, is *not* an ext. point of $C$, i.e  
    \[ \exists x_1 \neq x_2 \in C \text{ such that } x = \frac{x_1 + x_2}{2}. \]
  - If $x_1 \notin A$ or $x_2 \notin A$ then $\frac{1}{2} c^T (x_1 + x_2) > z = c^T x$ hence $x_1, x_2 \in A$ and thus $x$ is *not* an ext. point of $A$. 

look now for an extreme point of $A$. We use mainly $A \subset H_{c,z} \cap \mathbb{R}^m_+$

- if $A$ is a singleton, namely $A = \{y\}$, then $y$ is obviously extreme.
- if not, **narrow down recursively**:
  - $A^1 = \text{argmin}\{a_1 \mid a \in A\}$. 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:
    $$A^j = \text{argmin}\{a_j \mid a \in A^{j-1}\}.$$  

  - Since $A \subset \mathbb{R}^n$, 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 $A^k = \{a^k\}$.

- Suppose $\exists x^1 \neq x^2 \in A$ such that $a^k = \frac{x^1 + x^2}{2}$. In particular $\forall i \leq k$, $a_i^k = \frac{x_i^1 + x_i^2}{2}$.

- Since $a_1^k$ is an **infimum**, $x_i^1 = x_i^2 = a_i^k$ and $x^1, x^2 \in A^1$.

- By the same argument **applied recursively** we have that $x^1, x^2 \in A^j$ and finally $A^k$ which by construction is $\{a^k\}$, hence $x_1 = x_2 = a_k$, a contradiction, and $a^k$ is our extreme point.
Convex Hulls & Carathéodory’s Theorem
Convex combinations

Definition 6. Let \( \{x_1, x_2, \cdots, x_k\} \) 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 \( x = \sum_1^k \alpha_i x_i \) is called a **convex combination** of the points \( x_1, x_2, \cdots, x_k \).

Let’s illustrate this statement with a point \( x \) in a triangle \((x_1, x_2, x_3)\).

- Let \( y \) be the intersection of \((x_1, x)\) with \([x_2, x_3]\). \( y = px_2 + qx_3 \) with \( p = \frac{|x_2 - y|}{|x_3 - x_2|} \) and \( q = \frac{|x_3 - y|}{|x_3 - x_2|} \).

- On the other hand, \( x = lx_1 + ky \) with \( l = \frac{|x_1 - x|}{|x_1 - y|} \) and \( k = \frac{|y - x|}{|x_1 - y|} \).

- Finally \( x = lx_1 + pkx_2 + qkx_3 \), 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 $\iff$ all convex combinations

**Theorem 4.** The convex hull of a set of points $\{x_1, \cdots, x_k\}$ is the set of all convex combinations of $x_1, \cdots, x_k$.

**Proof.** Let $A = \{x \mid x = \sum_{i=1}^{k} \alpha_i x_i, \alpha_i \geq 0, \sum_{i=1}^{k} \alpha_i = 1\}$; $B = \langle\{x_1, \cdots, x_k\}\rangle$

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

$$
\sum_{i=1}^{k} (\lambda \alpha_i + (1 - \lambda)\beta_i) x_i \in A
$$

- $B \subseteq A$: $A$ is convex and contains each point $x_i$ since

$$
x_i = \sum_{j=1}^{k} \delta_{ij} x_j.
$$
Convex hull ⇔ all convex combinations

- $A \subseteq B$: by induction on $k$. If $k = 1$ then $B_1 = \langle \{x_1\} \rangle$ and $A_1 = \{x_1\}$. 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 $x \in A_k$ such that

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

If $x = x_k$ then trivially $x \in B_k$. If $x \neq x_k$ then $\alpha_k \neq 1$ and we have that

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

Consider $y = \sum_{i=1}^{k-1} \frac{\alpha_i}{1 - \alpha_k} x_i$. $y \in B_{k-1}$ by the induction hypothesis. Since $\{x_1, \ldots, x_{k-1}\} \subset \{x_1, \ldots, x_k\}$, $B_{k-1} \subseteq B_k$ by Lemma 2. Since $B_k$ is convex and both $y, x_k \in B_k$, so is $x = (1 - \alpha_k)y + \alpha_k x_k$. 

\[\blacksquare\]
Polytope, Polyhedrons

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

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

\[
P = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{r}_i^T \mathbf{x} \leq b_i, \ i = 1, \cdots, n \}\]

is called a **polyhedron**.

- A few comments:
  - **bounded polyhedron** ⇔ **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 $C = \{x_1, x_2, x_3, x_4, x_5\} \subset \mathbb{R}^2$ and its hull $\langle C \rangle$.

  $\begin{align*}
  y_1 & \text{ can be written as a convex combination of } x_1, x_2, x_3 \text{ (or } x_1, x_2, x_5); \\
  y_2 & \text{ can be written as a convex combination of } x_1, x_3, x_4; \\
  y_3 & \text{ can be written as a convex combination of } x_1, x_4, x_5;
  \end{align*}$

- For a set $C$ of 5 points in $\mathbb{R}^2$ there seems to be always a way to write a point $y \in \langle C \rangle$ as the convex combination of $2 + 1 = 3$ 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 $S \subset \mathbb{R}^n$. Then every point $x$ of $\langle S \rangle$ can be represented as a convex combination of $n + 1$ points from $S$,

$$ x = \alpha_1 x_1 + \cdots + \alpha_{n+1} x_{n+1}, \quad \sum_{i=1}^{n+1} \alpha_i = 1, \quad \alpha_i \geq 0. $$

**alternative formulation:**

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

- **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 $\mathbb{R}^m$. 
Proof.

- (⊇) is direct.
- (⊆) any $x \in \langle S \rangle$ can be written as a convex combination of $p$ points, $x = \alpha_1 x_1 + \cdots + \alpha_p x_p$. We can assume $\alpha_i > 0$ for $i = 1, \cdots, p$.
  - If $p < n+1$ then we add terms $0x_{p+1} + 0x_{p+2} + \cdots$ to get $n+1$ terms.
  - If $p > n+1$, we build a new combination with one term less:
    - let $A = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ 1 & 1 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n+1 \times p}$.
    - The key here is that since $p > n+1$ there exists a solution $\eta \in \mathbb{R}^m \neq 0$ to $A\eta = 0$.
    - By the last row of $A$, $\eta_1 + \eta_2 + \cdots + \eta_m = 0$, thus $\eta$ has both + and - coordinates.
    - Let $\tau = \min\left\{ \frac{\alpha_i}{\eta_i}, \eta_i > 0 \right\} = \frac{\alpha_i_0}{\eta_{i_0}}$.
    - Let $\tilde{\alpha}_i = \alpha_i - \tau \eta_i$. Hence $\tilde{\alpha}_i \geq 0$ and $\tilde{\alpha}_{i_0} = 0$.
      - $\tilde{\alpha}_1 + \cdots + \tilde{\alpha}_p = (\alpha_1 + \cdots + \alpha_p) - \tau(\eta_1 + \cdots + \eta_p) = 1$,
      - $\tilde{\alpha}_1 x_1 + \cdots + \tilde{\alpha}_p x_p = \alpha_1 x_1 + \cdots + \alpha_p x_p - \tau(\eta_1 x_1 + \cdots + \eta_p x_p) = x$.
    - Thus $x = \sum_{i \neq i_0} \alpha_i x_i$ of $p - 1$ points $\{x_i, i \neq i_0\}$.
    - Iterate this procedure until $x$ is a convex combin. of $n+1$ points of $S$. 
Next time

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