ORF 522

Linear Programming and Convex Analysis

A window on semidefinite programming

Marco Cuturi

So far

- Linear programming in $\mathbf{R}^n$
  - Simplex
  - Duality, dual simplex,
  - Structured constraints: network flows
  - Complexity: ellipsoid method
  - Efficiency: Interior Point Methods
  - Applications: OR, finance etc.
- A first generalization: Integer programs
  - cutting planes
  - branch& bound, branch & cut.
- Another?

Today

- finish this course with a **window** on **semi-definite programs**.
- A transition to ORF523.
- Semidefinite programming $=$ linear programming in the **cone of positive semidefinite matrices**.
- typically

$$\begin{array}{llc} \text{minimize} & \langle C, X\rangle & \\ \text{subject to} & \langle A_1, X\rangle & = b_1 \\ & \langle A_2, X\rangle & = b_2 \\ & \vdots & = \vdots \\ & \langle A_m, X\rangle & = b_m \\ & X \succeq \mathbf{0} & \end{array}$$

- Very very powerful tool. hot topic in last twenty years.
- Nesterov/Nemirovskii (1988) prove that IPM can be generalized to SDP's.
- After integer programs, a **further generalization** of LP's.
- Goal: focus on the **cone of semidefinite matrices** and its properties.

Faces and the Krein Milman Theorem

Reminder on Faces and Dimensions of Convex Sets

Definition 1. *Let C be a closed convex set. A set $F \subset C$ is called a* **face** *of C if there exists an affine hyperplane H which isolates C and such that $F = C \cap H$.*

Definition 2. *The* **dimension** *of a convex set $C \subset \mathbf{R}^d$ is the dimension of the smallest affine subspace that contains K*

- **remark**

1. A face K of dimension 0 is an **exposed point**.
2. A face K of dimension 1 is an **edge**.
3. A face K of dimension $d-2$ is called a **ridge**.
4. A face K of dimension $d-1$ is called a **facet**.

Combinations of Points

Given points $\mathbf{x}_1, \cdots, \mathbf{x}_m$, $\mathbf{x}$ is a

- **linear** combination if $\exists \lambda_1, \cdots, \lambda_m$ such that
- **affine** combination if $\exists \lambda_1, \cdots, \lambda_m$, $\sum_{i=1}^m \lambda_i = 1$ such that
- **convex** combination if $\exists \lambda_1, \cdots, \lambda_m \geq 0$, $\sum_{i=1}^m \lambda_i = 1$ such that
- **conic** combination if $\exists \lambda_1, \cdots, \lambda_m \geq 0$ such that

$$\mathbf{x} = \sum_{i=1}^{m} \lambda_i \mathbf{x}_i$$

Affine independence

Definition 3. *Points* $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ *in* $\mathbf{R}^n$ *are affinely independent (a.i.) if*

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k = 0, \quad a_1 + a_2 + \cdots + a_k = 0$$
$$\Downarrow$$
$$a_1 = a_2 = \cdots = a_k = 0$$

One can show that all the following statements are equivalent

(i) $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k \in \mathbf{R}^n$ are affinely independent

(ii) $\forall i \in \{1, \cdots, k\}$ vectors $\{\mathbf{x}_j - \mathbf{x}_i, j = 1, \cdots, k;\ j \neq i\}$ are l.i.

(iii) $\dim(\langle \mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k \rangle) = k - 1$

(iv) Every point of $\langle \mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k \rangle$ can be described as a **unique convex combination** of $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$. $\rightarrow$ "barycentric" coordinates.

(v) $\begin{pmatrix} \vdots & \vdots & \cdots & \vdots \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_k \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$ is invertible.

Interior points and Dimensionality

Theorem 1. *Let $C \subset \mathbf{R}^n$ be a convex set. If $\mathring{C} = \emptyset$ then there exists an affine subspace $L \subset \mathbf{R}^n$ such that $C \subset L$ and $\dim L < n$.*

- **Proof**: no $n+1$ affinely independent points in C.
  - if not, set $\Delta = \langle \mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{n+1} \rangle$ and we have $\Delta \subset C$.
  - Let $\mathbf{u} = \frac{1}{n+1} \sum_{i=1}^{n+1} \mathbf{x}_i$ be the barycenter of Δ.
  - For ε small enough $B_\varepsilon(\mathbf{u}) \subset \Delta$. Use invertibility of the matrix above.
  - Hence Δ has an interior point, C too, which is absurd.
- Let $k < n+1$ be the maximal number of affinely independent points in C.
- Then for each point $\mathbf{x}$ of C, there exists a collection of weights

$$\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k + \alpha \mathbf{x} = \mathbf{0}, \;\; \alpha_1 + \cdots + \alpha_k + \alpha = 0, \text{ with } \alpha \neq 0$$

- $\mathbf{x}$ can be expressed as the affine combination

$$\mathbf{x} = -\frac{\alpha_1}{\alpha}\mathbf{x}_1 - \frac{\alpha_1}{\alpha}\mathbf{x}_2 - \cdots - \frac{\alpha_k}{\alpha}\mathbf{x}_k$$

- Thus C lies in the affine hull L of $\mathbf{x}_1, \cdots, \mathbf{x}_k$ whose dimension is $k - 1 < n$.

Reminder on Faces

Lemma 1. *Let C be a closed convex set, and F a face of C such that $F = C \cap H \neq \emptyset$ where H is a supporting hyperplane of C. Then any extreme point of F is an extreme point of C.*

- F is a non-empty closed convex set.
- Let $H_{\mathbf{c},z}$ be a supporting hyperplane at $\mathbf{c} \in C$ and write $F = H_{\mathbf{c},z} \cap C$.
- **an extreme point of F is an extreme point of C**
  - suppose $\mathbf{x} \in F$, 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 $\mathbf{x}_1 \notin F$ **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 $\mathbf{x}_1, \mathbf{x}_2 \in F$ and thus $\mathbf{x}$ is **not** an ext. point of F.

Krein Milman

Theorem 2. *Let $C \subset \mathbf{R}^n$ be a compact convex set. Then C is the convex hull of the set of its extreme points, that is $C = \langle \mathbf{Ex}(C) \rangle$.*

- **Proof**: by induction on the dimension n of the ambient space.
- if $n = 0$ then C is a point and the result follows.
- suppose $n > 0$. if $\overset{\circ}{C} = \emptyset$ then it lies on a space of lower dimension and result is proved.
- suppose $n > 0$ and $\overset{\circ}{C} \neq \emptyset$. Let $\mathbf{u} \in C$.
  - if $\mathbf{u}$ is a boundary point, $\mathbf{u} \in \partial C$,
    - ▷ $\mathbf{u}$ belongs to a face F of C whose dimension is lower than n.
    - ▷ by recursion $\mathbf{u} \in \langle \mathbf{Ex}(F) \rangle$ and $\mathbf{Ex}(F) \subset \mathbf{Ex}(C)$.
  - if $\mathbf{u} \in \overset{\circ}{C}$,
    - ▷ let L be any arbitrary line (affine subspace of dim. 1) that contains $\mathbf{u}$.
    - ▷ $L \cap C = [\mathbf{a}, \mathbf{b}]$ where $\mathbf{a}, \mathbf{b} \in \partial C$.
    - ▷ $\mathbf{u}$ is a **convex** combination of $\mathbf{a}, \mathbf{b}$ which resp. belong to $\langle \mathbf{Ex}(C) \rangle$.

Existence of Extreme Points

Lemma 2. *Let $C \subset \mathbf{R}^n$ be a non-empty closed convex set which does not contain straight lines. Then C has an extreme point.*

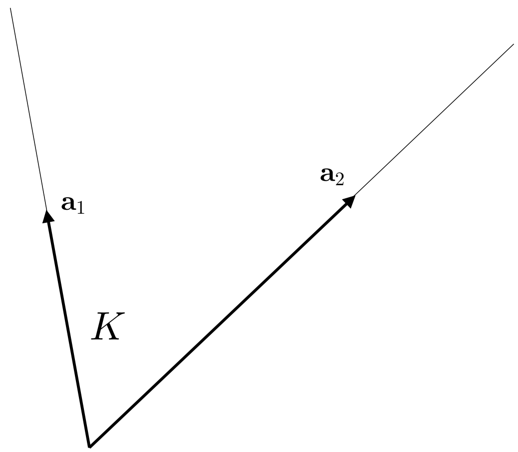
- **Proof**: similar to Krein-Milman..
- let's do it together.

- Direct corollary: a non-empty compact convex set has an extreme point.

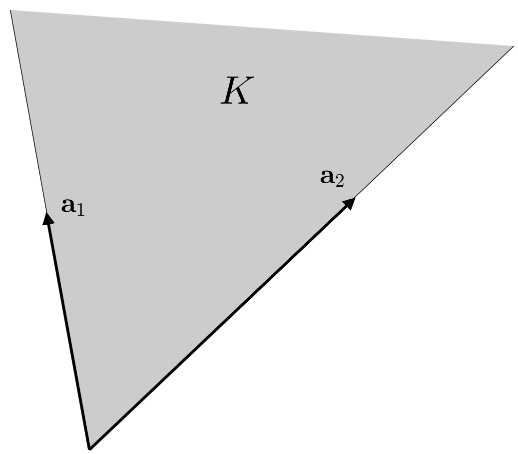
Convex Cones

Cones

- A set $K \subset \mathbf{R}^n$ is called a **cone** if $\forall \mathbf{x} \in K, \lambda \geq 0 \Rightarrow \lambda \mathbf{x} \in K$.



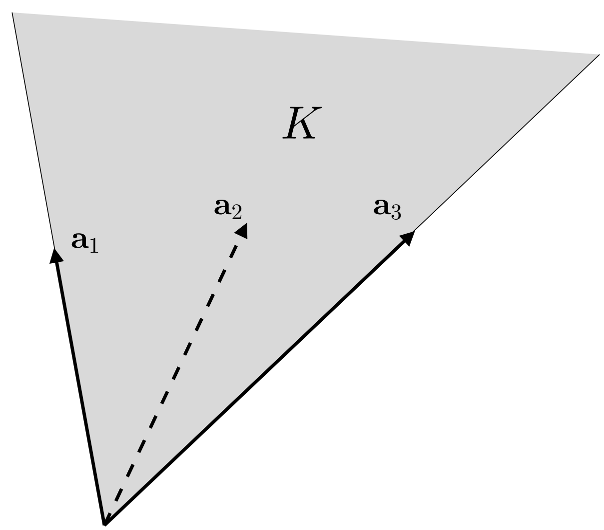
- Alternatively, a set K is a **convex cone** if $\forall \mathbf{x}, \mathbf{y} \in K, \alpha, \beta \geq 0 \Rightarrow \alpha \mathbf{x} + \beta \mathbf{y} \in K$.



Conic Hull & Rays

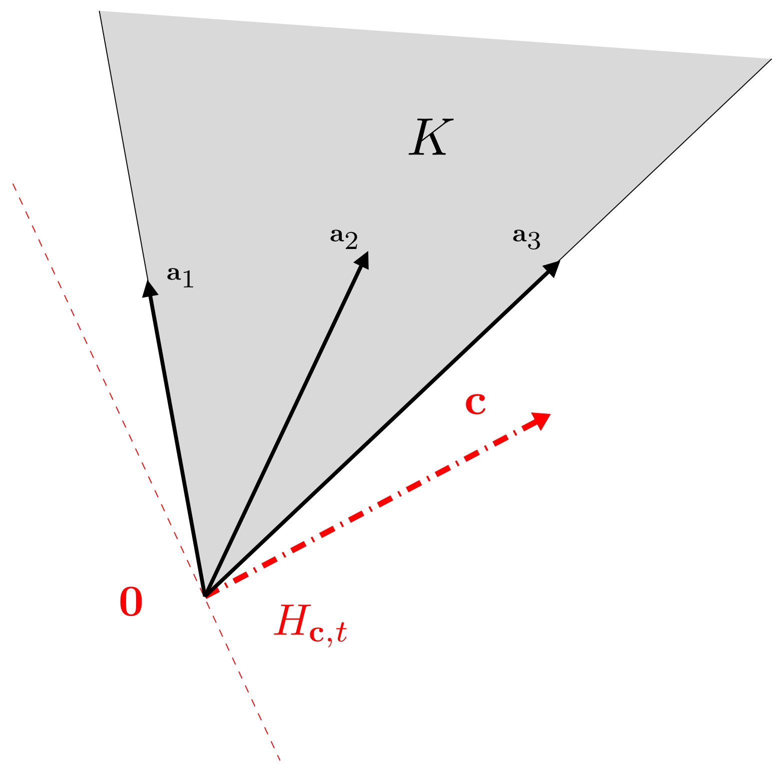
- The conic hull $\mathbf{Co}(S)$ is the set of all **conic** combinations of points taken in S.
- The conic hull $\mathbf{Co}(\mathbf{x})$ of a singleton $\{\mathbf{x}\}$ is called the **ray** spanned by $\mathbf{x}$.
- Let $K \subset \mathbf{R}^n$ be a cone and $K_1 \subset K$ a ray. K_1 is an **extreme** ray of K if for any $\mathbf{u} \in K_1$ and any $\mathbf{x}, \mathbf{y} \in K$

$$\mathbf{u} = \frac{\mathbf{x} + \mathbf{y}}{2} \Rightarrow \mathbf{x}, \mathbf{y} \in K_1$$



Isolating Hyperplanes and Cones

Lemma 3. *Let $K \subset \mathbf{R}^n$ be a cone and let $H \subset \mathbf{R}^n$ be an affine hyperplane isolating K and such that $K \cap H \neq \emptyset$. Then $\mathbf{0} \in H$.*



Proof: Let $\mathbf{y} \in K \cap H$. Assume $H = H_{\mathbf{c},t}$ and $K \subset \overline{H_+}$. By definition of K, $\mathbf{0} \in K$. Moreover, $\forall \mathbf{x} \in K$ $\mathbf{c}^T\mathbf{x} \geq \mathbf{c}^T\mathbf{y}$. Applying this to $\mathbf{0}$ we get $0 \geq t$. Suppose $t < 0$, that is $\mathbf{c}^T\mathbf{y} < 0$. Then for $\lambda > 1$, $\lambda\mathbf{c}^T\mathbf{y} < \mathbf{c}^T\mathbf{y}$ and thus $\mathbf{y}$ is in H_- while $\lambda\mathbf{y} \in K \subset \overline{H_+}$. Hence $t = 0$ and $\mathbf{0} \in H$.

Positive Definite Matrices

Symmetric Matrices

Definition 4. *A matrix* $A \in \mathbf{R}^{n\times n}$ *is called symmetric if* $A^T = A$.

- The space $\mathbf{Sym}_n$ of symmetric matrices is a vector space.
- It can be identified with $\mathbf{R}^{\frac{n(n+1)}{2}}$.
- A matrix $U \in \mathbf{R}^{n\times n}$ is **orthogonal** if $UU^T = I_n$ that is $U^T = U^{-1}$.
- For any matrix A in $\mathbf{Sym}_n$ there exists an orthogonal matrix U such that UAU^T is a **diagonal** matrix Δ
- This diagonal elements of Δ are the **eigenvalues** of A.
- The canonical scalar product of two symmetric matrices A and B is defined as

$$\langle A, B\rangle = \mathrm{tr}(AB) = \mathrm{tr}(BA).$$

- Note that for any orthogonal matrix U,

$$\langle A, B\rangle = \langle UAU^T, UBU^T\rangle$$

Positive Definite Matrices

Definition 5. *A matrix* $A \in \mathbf{Sym}_n$ *is positive definite (resp. semi-definite) if all its eigenvalues are positive (resp. nonnegative).*

- Alternative characterization: A is p.d. (resp. p.s.d.) if for all $\mathbf{x} \in \mathbf{R}^n$, $\mathbf{x} \neq \mathbf{0}$, $\mathbf{x}^T A\mathbf{x} > 0$ (resp. $\mathbf{x}^T A\mathbf{x} \geq 0$)
- We write $\mathbf{S}_n^+$ for the set of positive **semi-definite** matrices of size n
- For any matrix $A \in \mathbf{S}_n^+$ there exists U orthogonal such that $A = U\Delta U^T$ with Δ a **nonnegative** diagonal matrix.

A few properties (out of hundreds)

- Note that for all elements of a p.s.d. matrix A,

$$a_{ij}^2 \leq a_{ii} a_{jj}$$

why? *hint*: use $\Delta^{\frac{1}{2}}$ and the Cauchy-Schwartz inequality.

- Any diagonal entry of $A \in \mathbf{S}_n^+$ is non-negative. why? use $\mathbf{e}_i$.

- If $A \in \mathbf{S}_n^+$ and P any invertible matrix of $\mathbf{R}^{n \times n}$ then $PAP^{-1} \in \mathbf{S}_n^+$.

- If $A, B \in \mathbf{S}_n^+, \langle A, B \rangle \geq 0$. why? *hint*: decompose B into a sum of rank 1 matrices and compute $\langle A, B \rangle$

Interior of $\mathbf{S}_n^+$

Lemma 4. *A is an interior point of $\mathbf{S}_n^+$ iff A is p.d.*

- **Proof**: let A be in $\mathbf{S}_n^+$ and let $U = [\mathbf{u}_1, \cdots, \mathbf{u}_n]$ orthogonal such that $A = U\Delta U^T$.
- ($\Rightarrow$) suppose $\exists\varepsilon > 0$ such that $\forall M \in \mathbf{Sym}_n,\ \ \|A - M\|^2 < \varepsilon \Rightarrow M \in \mathbf{S}_n^+$.
  - Suppose $\exists j$ such that $A\mathbf{u}_j = 0$, i.e. A has a zero eigenvalue δ_j.
  - Let $A' = A + t\mathbf{u}_j\mathbf{u}_j^T$. $\|A - A'\| = t^2$ and $\mathbf{u}_j^T A'\mathbf{u}_j = t$.
  - taking $t < 0$ with $t^2 < \varepsilon$ we have $A' \in B_\varepsilon(A)$ but $\notin \mathbf{S}_n^+$.
- ($\Leftarrow$) suppose A is p.d. For all $j = 1, \cdots, n,\ \mathbf{u}_j^T A\mathbf{u}_j = \lambda_j > 0$.
  - For each $j = 1, \cdots, n,\ \exists\varepsilon_j$ such that $\forall M \in B_{\varepsilon_j}(A),\ \mathbf{u}_j^T M\mathbf{u}_j > 0$ by continuity of the function

$$\begin{array}{ccc} \mathbf{Sym}_n & \mapsto & \mathbf{R} \\ M & \rightarrow & \mathbf{u}_j^T M \mathbf{u}_j \end{array}.$$

  - Let $\varepsilon = \min\varepsilon_j$. Let $\mathbf{x} \in \mathbf{R}^n$ decomposed as $\sum_{i=1}^n x_i\mathbf{u}_i$ not be zero.
  - For $M \in B_\varepsilon(A), \mathbf{x}^T M\mathbf{x} = \sum_{i=1}^n x_i(\mathbf{u}_j^T M\mathbf{u}_j) \ \ > 0.$

Faces of S_n^+

Proposition 3. *Let $A \in \mathsf{S}_n^+$. Suppose that $\mathbf{Rank}(A) = r$. If $r = n$, A is an interior point of S_n^+. If $r < n$, A is an interior point of a face F of S_n^+, where $\dim(F) = r(r+1)/2$. There is a rank-preserving isometry identifying F with $\mathbf{Sym}_r$.*

- $r = n$ has just been solved.
- Suppose $\mathbf{Rank}(A) = r < n$. We build a suitable hyperplane $H \subset \mathbf{Sym}_n$ which contains A and isolates S_n^+.
  - Let $\lambda_1, \cdots, \lambda_r$ the non-zero eigenvalues of A.
  - Define U orthogonal such that $A = U\Delta U^T$ and $\Delta = \mathbf{diag}(\lambda_1, \cdots, \lambda_r, 0, \cdots, 0)$.
  - Let $C = \mathbf{diag}(0, \cdots, 0, 1, \cdots, 1)$ be the diagonal matrix of r zeroes and $n - r$ ones.
  - Let $Q = UCU^T$. Obviously $Q \in \mathsf{S}_n^+$ and $\langle A, Q\rangle = 0$.
  - Furthermore, $\forall Y \in \mathsf{S}_n^+, \langle Y, Q\rangle = \langle U^T Y U, C\rangle \geq 0$.
  - Therefore $H = \{X \in \mathbf{Sym}_n | \langle Q, X\rangle = 0\}$ isolates S_n^+ and contains A.
  - Set $F = \mathsf{S}_n^+ \cap H$. The map $\varphi : X \to Y = U^T X U$ maps Q onto C and A onto D.
  - $\varphi(F) = F' = \{Y \in \mathbf{Sym}_n | \langle C, Y\rangle = 0\}$. Let $Y \in F'$.

- By nonnegativity of its diagonals, $y_{jj} = 0$ for $j \geq r+1$. Y must thus have the following block structure

$$Y = \begin{bmatrix} W_{r\times r} & \mathbf{0}_{r\times n-r} \\ \mathbf{0}_{n-r\times r} & \mathbf{0}_{n-r\times n-r} \end{bmatrix},$$

  with $W_{r\times r} \in \mathbf{S}_r^+$
- Hence the face F' can be identified with $\mathbf{S}_r^+$ and $\mathbf{S}_r^+$ contains D in its interior.
- Since $\varphi^{-1} : Y \mapsto X = UYU^T$ is a non-degenerate linear transformation, which maps D to A and F' to F,
- we have $\dim(F) = r(r+1)/2$ and F contains A in its interior.

Next time

- Linear equation in $\mathbf{S}_n^+$.