ORF 522

Linear Programming and Convex Analysis

Linear Equations in Semidefinite Matrices

Marco Cuturi

Reminder

- Convexity
  - Affine independence.
  - Faces, Dimension, Interior.
  - Krein-Milman.
  - No straight lines in a closed convex-set $\Rightarrow \exists$ extreme point.
- Positive semidefinite matrices
  - Identify $\mathbf{Sym}_n$ with $\mathbb{R}^{n(n+1)/2}$ although Frobenius dot-product slightly different.
  - $\mathbf{S}_n^+ =$ subset of matrices of $\mathbf{Sym}_n$ with nonnegative eigenvalues
  - A interior point of $\mathbf{S}_n^+ \leftrightarrow A$ is positive definite.

Today

- A few more results on $\mathbf{S}_n^+$
- "Linear" programming $\rightarrow$ study linear equations in $\mathbf{S}_n^+$.
- A simple application in embeddings.

Further results on $\mathbf{S}_n^+$

$\mathbf{S}_n^+$ is a closed-convex cone with no straight lines

Proposition 1. *The set of positive semidefinite matrices is a closed convex cone which does not contain straight lines.*

- $A, B \in \mathbf{S}_n^+$, $\alpha, \beta \geq 0$ then $\forall \mathbf{x} \in \mathbf{R}^n$ $\mathbf{x}^T(\alpha A + \beta B)\mathbf{x} = \alpha \mathbf{x}^T A \mathbf{x} + \beta \mathbf{x}^T B \mathbf{x} \geq 0$.
- closed: convergence of nonnegative eigenvalues.
- straight lines: for any two matrices A, B, any line $\{B + \lambda A, \lambda \in \mathbf{R}\}$, cannot be entirely in $\mathbf{S}_n^+$.

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

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

- $r = n$ proved in previous theorem.
- Suppose $\mathbf{Rank}(A) = r < n$. We build a suitable hyperplane $H \subset \mathbf{Sym}_n$ which contains A and isolates $\mathbf{S}_n^+$.
  - Let $\lambda_1, \cdots, \lambda_r$ the non-zero eigenvalues of A.
  - Define U orthogonal such that $A = UDU^T$ and $D = \mathbf{diag}(\lambda_1, \cdots, \lambda_r, 0, \cdots, 0)$.
  - 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 $Q \in \mathbf{S}_n^+$ and $\langle A, Q \rangle = 0$.
  - Furthermore, $\forall Y \in \mathbf{S}_n^+, \langle Y, Q \rangle = \langle U^T Y U, C \rangle \geq 0$.
  - Therefore $H = \{X \in \mathbf{Sym}_n | \langle Q, X \rangle = 0\}$ isolates $\mathbf{S}_n^+$ and contains A.
  - Set $F = \mathbf{S}_n^+ \cap H$. 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 diagonal elements, $y_{jj} = 0$ 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 $\mathbf{S}_r^+$
- $\mathbf{S}_r^+$ 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.

Linear Equations in $\mathbf{S}_n^+$

Linear Equations in $\mathbf{S}_n^+$

Proposition 3. *Let us fix k matrices $A_1, \cdots, A_k$ matrices in $\mathbf{Sym}_n$ and k real numbers $\alpha_1, \cdots, \alpha_k$. If there exists a matrix $X \in \mathbf{S}_n^+$ such that*

$$\langle A_i, X \rangle = \alpha_i, i = 1, \cdots, k$$

then there exists a matrix $X_0 \in \mathbf{S}_n^+$ such that

$$\langle A_i, X_0 \rangle = \alpha_i, i = 1, \cdots, k$$

and additionally such that $\mathbf{Rank}(X_0) \leq \lfloor \frac{\sqrt{8k+1}-1}{2} \rfloor$.

is equivalent to

Proposition 4. *Let $\mathcal{A} \subset \mathbf{Sym}_n$ be an affine subspace such that the intersection $\mathbf{S}_n^+ \cap \mathcal{A}$ is non-empty. Suppose* $\dim(\mathcal{A}) > n(n+1)/2 - (r+1)(r+2)/2$ *for some non-negative integer r. Then there is a matrix X in $\mathbf{S}_n^+ \cap \mathcal{A}$ such that* $\mathbf{Rank}(X) \leq r$

Linear Equations in $\mathbf{S}_n^+$

- **Proof of equivalence**:
  - Let $A = \{X \in \mathbf{Sym}_n : \langle A_i, X\rangle = \alpha_i,\ i = 1, \cdots, k\}$.
  - Then $\dim(A) \geq n(n+1)/2 - k$.
  - Moreover $k < (r+2)(r+1)/2$ iff $r \leq \lfloor \frac{\sqrt{8k+1}-1}{2} \rfloor$
  - why? if $x = (y+2)(y+1)/2$ then $y = \frac{\pm\sqrt{8x+1}-1}{2}$.

- **Proof**: we prove the second proposition.
  - Let $\mathcal{K} = \mathcal{A} \cap \mathbf{S}_n^+$. The set $\mathcal{K}$ is non empty, closed and does not contain straight lines $\rightarrow$ it **contains an extreme point** X_0.
  - Suppose $\mathbf{Rank}(X_0) = m$. Thus by Proposition 2 X_0 must be an **interior point** of a face F of $\mathbf{S}_n^+$, embedded in $\mathbf{S}_m^+$ of dimension $m(m+1)/2$.
  - $X_0 \in \mathcal{K}$. X_0 is an interior point of the intersection $F \cap \mathcal{A}$.
  - Since X_0 is an extreme point, we must have $\dim(F \cap \mathcal{A}) = 0$ (why?).
  - This implies $\dim(F) + \dim(\mathcal{A}) < n(n+1)/2$ hence $\dim(\mathcal{A}) > n(n+1)/2 - (m+1)(m+2)/2$.
  - Hence $m \leq r$ and $\exists X_0 \in \mathcal{A} \cap \mathbf{S}_n^+$ such that $\mathbf{Rank}(X_0) \leq r$.

Some further comments

- In the real case, solutions of

$$A\mathbf{x} = \mathbf{b}, \;\; A \in \mathbf{R}^{m \times n},$$

can be found such that $\mathbf{x}$ has at least $n - m$ zero components or up to n non-zero components.

- In the psd case, representing $X \in \mathbf{S}_n^+$ by a vector $X(:)$ of size $n(n+1)/2$,

$$\mathbf{A}X(:) = \mathbf{b}, \;\; \mathbf{A} \in \mathbf{R}^{m \times n\frac{n+1}{2}}$$

we have the result that a solution X with up to $\lfloor \frac{\sqrt{8k+1}-1}{2} \rfloor$ non-zero eigenvalues can be obtained.

- Some work on extensions of the simplex show that "extreme points" on the set $\mathcal{A}$ are low (r such that $r(r+1)/2 \leq m$ more precisely) rank matrices (Pataki, 1996). generalization is not straightforward however.

A more advanced result

Proposition 5. *Let us fix k matrices $A_1, \cdots, A_k$ matrices in $\mathbf{Sym}_n$, where $k = (r+1)(r+2)/2$ with $r > 0$ and $n \geq r+2$, and k real numbers $\alpha_1, \cdots, \alpha_k$. If there exists a matrix $X \in \mathsf{S}_n^+$ such that*

$$\langle A_i, X\rangle = \alpha_i, i = 1, \cdots, k$$

and the set of ***all solutions to these equations is bounded****, then there exists a matrix $X_0 \in \mathsf{S}_n^+$ such that*

$$\langle A_i, X\rangle = \alpha_i, i = 1, \cdots, k$$

and additionally such that $\mathbf{Rank}(X_0) \leq r$.

is equivalent to

Proposition 6. *Let $\mathcal{A} \subset \mathbf{Sym}_n$ be an affine subspace such that the intersection $\mathsf{S}_n^+ \cap \mathcal{A}$ is non-empty* ***and bounded****. Suppose*

$$\dim(\mathcal{A}) = n(n+1)/2 - (r+1)(r+2)/2$$

for some positive integer r and $n \geq r+2$. Then there is a matrix X in $\mathsf{S}_n^+ \cap \mathcal{A}$ such that $\mathbf{Rank}(X) \leq r$

What is the difference

- Proof is quite involved (a few pages, uses topology)
- In practice, for a number of constraints k if the set of solutions is not empty, the minimal rank solution is of rank r,
  - $k = 3$, $r \leq 1$,
  - $k = 6$, $r \leq 2$,
  - $k = 10$, $r \leq 3$
- compared to the bounds of Proposition 3:
  - $k = 3$, $r \leq 2$
  - $k = 6$, $r \leq \frac{\sqrt{8 \cdot 6 + 1} - 1}{2} = 3$
  - $k = 10$, $r \leq \frac{\sqrt{8 \cdot 10 + 1} - 1}{2} = 4$
- Existence theorems only.
- Recovering a solution of low rank from an arbitrary solution requires iterative algorithms

Approximation

Proposition 7. *Let us fix k matrices $A_1, \cdots, A_k$ matrices in $\mathbf{S}_n^+$, k nonnegative numbers $\alpha_1, \cdots, \alpha_k$ and a number $0 < \varepsilon < 1$. If there exists a matrix $X \in \mathbf{S}_n^+$ such that*

$$\langle A_i, X \rangle = \alpha_i, i = 1, \cdots, k$$

then, letting m be a positive integer such that

$$m \geq \frac{8}{\varepsilon^2} \ln(4k),$$

there exists a matrix $X_0 \in \mathbf{S}_n^+$ such that

$$\alpha_i(1 - \varepsilon) \leq \langle A_i, X_0 \rangle \leq \alpha_i(1 + \varepsilon), i = 1, \cdots, k$$

and additionally such that $\mathbf{Rank}(X_0) \leq m$.

- No proof, but look at the improvement with approximation: from $\mathbf{Rank}(X_0) = O(\sqrt{k})$ to $\mathbf{Rank}(X_0) = O(\ln k)$.
- These results are in Barvinok (2002)

An application: Graph Realizability

Gram matrices

- For $\mathbf{x}_1, \cdots, \mathbf{x}_m$ vectors in $\mathbf{R}^n$, the matrix

$$K = [k_{ij}]_{1 \leq i,j \leq n},$$

defined as

$$k_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$$

is called the **Gram** matrix of vectors $\mathbf{x}_1, \cdots, \mathbf{x}_m$.

- $\mathbf{Rank}(K) = \dim(\mathrm{span}\{\mathbf{x}_1, \cdots, \mathbf{x}_m\}) \leq \min(n, m)$ why?
  - Set $X = [\mathbf{x}_1, \cdots, \mathbf{x}_m] \in \mathbf{R}^{n \times m}$.
  - Then $K = X^T X \in \mathbf{R}^{m \times m}$
  - Can show that $\ker(K) = \ker(X)$

- Conversely, can prove that if $K \in \mathbf{S}_n^+$ and $\mathbf{Rank}(K) \leq r$ then K is the gram matrix of vectors in $\mathbf{R}^r$

Graph Realization Problem

- Suppose we are given an undirected weighted graph $\mathcal{G} = (\mathcal{N}, \mathcal{E}, \rho)$ where
  - $\mathcal{N}$ is the set of nodes $(v_1, \cdots, v_n)$
  - $\mathcal{E}$ the set of edges
  - ρ is a family of weights indexed by the edges $\rho_e \in \mathbf{R}$ for every $e \in \mathcal{E}$.

  Definition 1. *A weighted graph $\mathcal{G}(\mathcal{N}, \mathcal{E}, \rho)$ is d-realizable if there exists a way to associate to each node $v_1, \cdots, v_n$ vectors $\mathbf{v}_1, \cdots, \mathbf{v}_n \in \mathbf{R}^d$ respectively such that $\|\mathbf{v}_i - \mathbf{v}_j\| = \rho_{\{i,j\}}$.*

- A weighted graph is realizable if it is d-realizable for an certain dimension d.

Realizability

- An important problem: 3-realizability,
  - molecular conformation: atoms, distances imposed by physical lows, which configurations are possible?
  - industry: in which configurations can a few joints connected by rigid links move?
  - sensor network configuration
- Existence of low-realizability given distances is also used in data-visualization (low dimensional embeddings)

A straightforward reformulation

- Let $\mathbf{v}_1, \cdots, \mathbf{v}_n \in \mathbf{R}^d$ be a realization of the graph in $\mathbf{R}^d$
- let $K = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$ be the gram matrix of $\mathbf{v}_1, \cdots, \mathbf{v}_n$.
- $K \in \mathbf{S}_n^+$. why?
- For any edge $\{i, j\}$,

$$\rho_{\{i,j\}}^2 = \|\mathbf{v}_i - \mathbf{v}_j\|^2 = \|\mathbf{v}_i\|^2 + \|\mathbf{v}_j\|^2 - 2\langle \mathbf{v}_i, \mathbf{v}_j \rangle = k_{jj} + k_{ii} - 2k_{ij}.$$

- Can be interpreted as $|\mathcal{E}|$ constraints
- The d-realizability problem is equivalent to looking for a matrix $X \in \mathbf{S}_n^+$ with the additional constraint that $\mathbf{Rank}(X) \leq d$.

Realizability and d-realizability

Proposition 8. *Suppose that* $|\mathcal{E}| \leq (d+1)(d+2)/2$. *Then* $\mathcal{G}$ *is d-realizable if and only if it is realizable. In particular if* $k \leq 9$ *then the graph is realizable iff it is 3-realizable.*

- **Proof**: follows from of proposition 3.
- Comment: realizability only depends on the number of edges, not nodes.
- Edges for which such a constraint is not given can be freely set.

With approximations: Johnson-Lindenstrauss Lemma

- Proof uses the approximation result of Proposition 7 we discussed before

  Proposition 9. *Suppose that a graph* $\mathcal{G}$ *with* k *edges is realizable. Then for any* $0 < \varepsilon < 1$ *and any* $m \geq \frac{8}{\varepsilon^2}\ln(4k)$ *one can place the nodes* $v_1, \cdots, v_n$ *on points* $\mathbf{v}_1, \cdots, \mathbf{v}_n$ *in* $\mathbf{R}^m$ *so that*

$$\rho_{\{i,j\}}(1-\varepsilon) \leq \|\mathbf{v}_i - \mathbf{v}_j\|^2 \leq \rho_{\{i,j\}}(1+\varepsilon), \ \ \{i,j\} \in \mathcal{E}.$$

- **Proof**
  - Define a constraint matrix $A_{i,j}$ for each edge's constraint.
  - Show that each constraint matrix $A_{i,j}$ is $\mathbf{S}_n^+$
  - Since $\mathcal{G}$ is realizable, we can use the approximation result of Proposition 7 directly.

- Existence result, often seen as an objective for dimensionality reduction algorithms

Final Exam

Description & Questions

- 3 hours, Thu. 14th, 14:00$\rightarrow$ 17:00, room 001 downstairs.
  - ≈ 1 hour for short questions / multiple choice questions
  - 2 small exercises to check your understanding of the lectures.
  - 1 problem to see how you can generalize from lectures.
- Each part graded proportionally.
- A letter format cheat sheet is allowed, nothing else.