Introduction to Semidefinite Programming I: Basic properties and variations on the Goemans-Williamson approximation algorithm for max-cut

MFO seminar on Semidefinite Programming

Monique Laurent - CWI, Amsterdam, and Tilburg University

May 30, 2010

Positive semidefinite matrices

Definition: For a symmetric $n \times n$ matrix X, the following conditions are equivalent:

1. X is *positive semidefinite* (written $X \succeq 0$) if all eigenvalues of X are nonnegative.
2. $u^T X u \geq 0$ for all $u \in \mathbb{R}^n$.
3. $X = UU^T$ for some matrix $U \in \mathbb{R}^{n \times p}$.
4. For some vectors $v_1, \ldots, v_n \in \mathbb{R}^p$, $X_{ij} = v_i^T v_j$ ($i, j \in [n]$). Say that X is the *Gram matrix* of the v_i's.
5. All principal minors of X are nonnegative.

Definition: X is *positive definite* (written $X \succ 0$) if all eigenvalues of X are positive.

Notation

- $\mathcal{S}_n$: the space of $n \times n$ symmetric matrices.
- $\mathcal{S}_n^+$: the cone of positive semidefinite matrices.
- $\mathcal{S}_n^{++}$: the cone of positive definite matrices.

$\rightsquigarrow$ $\mathcal{S}_n^{++}$ is the interior of the cone $\mathcal{S}_n^+$.

Trace inner product on $\mathcal{S}_n$:

$$A \cdot B = Tr(A^T B) = \sum_{i,j=1}^{n} A_{ij} B_{ij}$$

The PSD cone is self-dual: For $X \in \mathcal{S}_n$,

$$A \in \mathcal{S}_n^+ \iff A \cdot B \geq 0 \ \ \forall B \in \mathcal{S}_n^+$$

Primal/dual semidefinite programs

Given matrices $C, A_1, \ldots, A_m \in \mathcal{S}_n$ and a vector $b \in \mathbb{R}^m$

Primal SDP:

$$p^* := \max_X \; C \cdot X \; \text{ such that } \; A_j \cdot X = b_j \; (j = 1, \ldots, m), \; X \succeq 0$$

Dual SDP:

$$d^* := \min_y \; b \cdot y \; \text{ such that } \; \sum_{j=1}^m y_j A_j - C \succeq 0$$

Weak duality: $p^* \leq d^*$

Pf: If X is *primal feasible* and y is *dual feasible*, then

$$0 \leq \left(\sum_{j=1}^{m} y_j A_j - C \right) \cdot X = \sum_j y_j (A_j \cdot X) - C \cdot X = b \cdot y - C \cdot X$$

Analogy between LP and SDP

Given vectors $c, a_1, \ldots, a_m \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$

Primal/dual LP:

$$\max_x \; c \cdot x \;\text{ such that }\; a_j \cdot x = b_j \; (\forall j \leq m), \; x \in \mathbb{R}^n_+$$

$$\min_y \; b \cdot y \;\text{ such that }\; \sum_{j=1}^m y_j a_j - c \geq 0$$

Primal SDP:

$$\max_X \; C \cdot X \;\text{ such that }\; A_j \cdot X = b_j \; (j = 1, \ldots, m), \; X \in \mathcal{S}_n^+$$

- SDP is the analogue of LP, replacing $\mathbb{R}^n_+$ by $\mathcal{S}_n^+$.
- Get LP when C, A_j are diagonal matrices.

Strong duality: $p^* = d^*$?

Strong duality holds for LP, but we need some regularity condition (e.g., *Slater condition*) to have strong duality for SDP !

Primal (P) / Dual (D) SDP's: $p^* \leq d^*$

(P) $p^* = \sup \ C \cdot X \ \text{ s.t. } \ A_j \cdot X = b_j \ (j = 1, \ldots, m), \ X \succeq 0$

(D) $d^* = \inf \ b \cdot y \ \text{ s.t. } \ \sum_{j=1}^m y_j A_j - C \succeq 0$

Strong duality Theorem:

1. If (P) is *strictly feasible* ($\exists X \succ 0$ feasible for (P)) and bounded ($p^* < \infty$), then $p^* = d^*$ and (D) attains its infimum.
2. If (D) is *strictly feasible* ($\exists y$ with $\sum_j y_j A_j - C \succ 0$) and bounded ($d^* > -\infty$), then $p^* = d^*$ and (P) attains its supremum.

Proof of 2. Assume $d^* \in \mathbb{R}$ and $\sum_j \tilde{y}_j A_j - C \succ 0 \;\; \exists \tilde{y}$

$$p^* = \max_{X \succeq 0} \; C \cdot X \;\; A_j \cdot X = b_j \quad \overset{?}{\geq} \quad d^* = \inf_y \; b \cdot y \;\; \sum_j y_j A_j - C \succeq 0$$

Goal: There exists X feasible for (P) with $C \cdot X \geq d^*$.

WMA $b \neq 0$ (else, $b = 0$ implies $d^* = 0$ and choose $X = 0$). Set

$$\mathcal{M} := \left\{ \sum_j y_j A_j - C \mid y \in \mathbb{R}^m, \; b \cdot y \leq d^* \right\}.$$

Fact: $\mathcal{M} \cap \mathcal{S}_n^{++} = \emptyset$.

Pf: Otherwise, let y for which $b \cdot y \leq d^*$ and $\sum_j y_j A_j - C \succ 0$.
Then one can find y' with $b \cdot y' < b \cdot y \leq d^*$ and $\sum_j y_j' A_j - C \succ 0$.
$\rightsquigarrow$ This contradicts the minimality of d^*.

Sketch of proof for 2. (continued)

As $\mathcal{M} \cap \mathcal{S}_n^{++} = \emptyset$, there is a hyperplane separating $\mathcal{M}$ and $\mathcal{S}_n^{++}$.

That is, there exists $Z \succeq 0$ non-zero with $Z \cdot Y \leq 0 \ \forall Y \in \mathcal{M}$, i.e.,

$$b \cdot y \leq d^* \Longrightarrow Z \cdot (\sum_j y_j A_j - C) \leq 0$$

By Farkas' lemma, there exists $\mu \in \mathbb{R}_+$ for which

$$(Z \cdot A_j)_j = \mu b \ \text{ and } \ \mu d^* \leq Z \cdot C$$

If $\mu = 0$, then $0 \geq \underbrace{Z}_{\succeq 0} \cdot \underbrace{(\sum_j \tilde{y}_j A_j - C)}_{\succ 0} > 0$, a contradiction.

Hence $\mu > 0$ and Z/μ is feasible for (P) with $C \cdot (Z/\mu) \geq d^*$.
QED.

An example with duality gap

$$p^* = \min\ x_{12}\ \text{ s.t. } \begin{pmatrix} 0 & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & 1+x_{12} \end{pmatrix} \succeq 0$$

$$\begin{aligned} = \min \tfrac{1}{2}E_{12}\cdot X \ \text{ s.t. } \quad & E_{11}\cdot X = 0 && \rightsquigarrow a \\ & E_{13}\cdot X = 0 && \rightsquigarrow b \\ & E_{23}\cdot X = 0 && \rightsquigarrow c \\ & (E_{33} - \tfrac{1}{2}E_{12})\cdot X = 1 && \rightsquigarrow y \\ & X \succeq 0 \end{aligned}$$

$$d^* = \max\ y \text{ s.t. } \tfrac{1}{2}E_{12} - aE_{11} - bE_{13} - cE_{23} - y(E_{33} - \tfrac{1}{2}E_{12}) \succeq 0$$

$$= \max\ y \text{ s.t. } \begin{pmatrix} -a & \frac{y+1}{2} & -b \\ \frac{y+1}{2} & 0 & -c \\ -b & -c & -y \end{pmatrix} \succeq 0$$

Thus, $p^* = 0$, $d^* = -1 \rightsquigarrow$ non-zero duality gap

Complexity

Recall: An LP with rational data has a *rational* optimum solution whose *bit size is polynomially bounded* in terms of the bit length of the input data.

Not true for SDP:

- $\sqrt{2} = \max\ x \ \text{ s.t. } \begin{pmatrix} 1 & x \\ x & 2 \end{pmatrix} \succeq 0$
- Any solution to
$$\begin{pmatrix} x_1 - 2 & 0 \\ 0 & 1 \end{pmatrix} \succeq 0, \begin{pmatrix} x_2 & x_1 \\ x_1 & 1 \end{pmatrix} \succeq 0, \ldots, \begin{pmatrix} x_n & x_{n-1} \\ x_{n-1} & 1 \end{pmatrix} \succeq 0$$
satisfies $x_1 \geq 2^{2^{n-1}}$.

Complexity (continued)

Theorem: SDP can be solved in polynomial time to an arbitrary prescribed precision. [Assuming certain technical conditions hold.]

- Theoretically: Use the ellipsoid method [since checking whether $X \succeq 0$ is in P, e.g. with Gaussian elimination]
- Practically: Use e.g. interior-point algorithms.

More precisely: Let K denote the feasible region of the SDP. Assume we know $R \in \mathbb{N}$ s.t. $\exists X \in K$ with $\|X\| \leq R$ if $K \neq \emptyset$.

Given $\epsilon > 0$, the ellipsoid based algorithm, either finds X^* at distance at most ϵ from K such that $C \cdot X^* \geq C \cdot X - \epsilon \;\; \forall X \in K$ at distance at least ϵ from the border, or claims: there is no such X.

The rumnning time is polynomial in n, m, the bit size of A_j, C, b, $\log R$, and $\log(1/\epsilon)$.

Feasibility of SDP

Feasibility SDP problem (F): Given integer $A_0, A_j \in \mathcal{S}_n$, decide whether there exists $x \in \mathbb{R}^m$ s.t. $A_0 + \sum_{j=1}^m x_j A_j \succeq 0$?

- (F) $\in$ NP $\iff$ (F) $\in$ co-NP. [Ramana 97]
- (F) $\in$ P for **fixed** n or m. [Porkolab-Khachiyan 97]
- Testing existence of a *rational* solution is in P for **fixed dimension** m.

  [Porkolab-Khachiyan 97]

$\rightsquigarrow$ More on complexity and algorithms for SDP in other lectures.

Use SDP to express convex quadratic constraints

Consider the convex quadratic constraint:

$$x^T A x \leq b^T x + c$$

where $A \succeq 0$.

- Write $A = B^T B$ for some $B \in \mathbb{R}^{p \times n}$.
- **Then:** $x^T A x \leq b^T x + c \iff \begin{pmatrix} I_p & Bx \\ x^T B^T & b^T x + c \end{pmatrix} \succeq 0$

$\rightsquigarrow$ Use **Schur complement:** Given $C \succ 0$,

$$\begin{pmatrix} C & B \\ B^T & A \end{pmatrix} \succeq 0 \iff A - B^T C^{-1} B \succeq 0$$

The S-lemma [Yakubovich 1971]

Consider the quadratic polynomials:

$$f(x) = x^T Ax + 2a^T x + \alpha = (1\ x^T) \begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}$$

$$g(x) = x^T Bx + 2b^T x + \beta = (1\ x^T) \begin{pmatrix} \beta & b^T \\ b & B \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}$$

Question: Characterize when

$$(^*)\ f(x) \geq 0 \Longrightarrow g(x) \geq 0$$

Answer: Assume $f(x) > 0$ for some x. Then, (*) holds IFF

$$\begin{pmatrix} \beta & b^T \\ b & B \end{pmatrix} - \lambda \begin{pmatrix} \alpha & a^T \\ a & A \end{pmatrix} \succeq 0 \ \text{ for some } \lambda \geq 0$$

Testing sums of squares of polynomials

Question: How to check whether a polynomial

$p(x) = \sum_{\alpha \in \mathbb{N}^n | |\alpha| \leq 2d} p_\alpha x^\alpha$ can be written as a **sum of squares:**

$p(x) \stackrel{?}{=} \sum_{j=1}^m (u_j(x))^2$ for some polynomials u_j?

$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$

Answer: Use semidefinite programming:

- Write $u_j(x) = (a_j)^T [x]_d$ $\quad [x]_d = (x^\alpha)_{|\alpha| \leq d}$
- $\sum_j (u_j(x))^2 = ([x]_d)^T \underbrace{\left(\sum_j a_j a_j^T\right)}_{\rightsquigarrow A \succeq 0} [x]_d$

$\rightsquigarrow$ Test feasibility of SDP:

$$\sum_{\substack{\beta,\gamma : |\beta|, |\gamma| \leq d \\ \beta+\gamma=\alpha}} A_{\beta,\gamma} = p_\alpha \ (|\alpha| \leq 2d), \ A \succeq 0$$

Two milestone applications of SDP to combinatorial optimization

- Approximate maximum stable sets and minimum vertex coloring with the theta number.

  Work of Lovász [1979], Grötschel-Lovász-Schrijver [1981]

- (First non-trivial) 0.878-approximation algorithm for max-cut of Goemans-Williamson [1995]

The theta number

$G = (V, E)$ graph.

$S \subseteq V$ **stable set** if S contains no edge.

- $\alpha(G)$:= maximum size of a stable set $\rightsquigarrow$ stability number
- $\chi(G)$:= minimum number of colors needed to color the vertices so that adjacent vertices receive distinct colors.

  $\rightsquigarrow$ Computing $\alpha(G), \chi(G)$ is an NP-hard problem.
- **The theta number of Lovász [1979]:**

$$\vartheta(G) := \max J{\cdot}X \quad \text{s.t.} \quad Tr(X) = 1,\ X_{ij} = 0\ (ij \in E),\ X \succeq 0$$

- **Lovász 'sandwich' theorem:** $\alpha(G) \leq \vartheta(G) \leq \chi(\bar{G})$.

$\rightsquigarrow$ Can compute $\alpha(G)$, $\chi(\bar{G})$ via SDP for graphs with $\alpha(G) = \chi(\bar{G})$.

Maximum cuts in graphs

$G = (V, E)$, $n = |V|$, $w = (w_e)_{e \in E}$ edge weights.

$S \subseteq V \rightsquigarrow$ **cut** $\delta(S) :=$ all edges **cut** by the partition $(S, V \setminus S)$.

Max-Cut problem: Find a cut of maximum weight. $\rightsquigarrow \mathrm{mc}(G)$

- Max-Cut is NP-hard.
  No $(16/17 + \epsilon)$-approximation algorithm unless P=NP.
- Max-Cut is in P for graphs with no K_5 minor, since it can be computed with the LP: [Barahona-Mahjoub 86]

  $\max\ w^T x$ s.t. $x_{ij} - x_{ik} - x_{jk} \leq 0,\ x_{ij} + x_{ik} + x_{jk} \leq 2\ \forall i, j, k \in V$

An easy $1/2$-**approximation algorithm** for $w \geq 0$: Consider the random partition $(S, V \setminus S)$, where $i \in S$ with prob. $1/2$:

$$E(w(S)) = w(E)/2 \geq \mathrm{mc}(G)/2$$

Goemans-Williamson approximation algorithm for Max-Cut

- Encode a partition $(S, V \setminus S)$ by a vector $x \in \{\pm 1\}^n$.

$\rightsquigarrow$ Encode the cut $\delta(S)$ by the matrix $X = xx^T$.

$\rightsquigarrow$ **Reformulate Max-Cut:**

$$\max \frac{1}{2} \sum_{ij \in E} w_{ij}(1 - x_i x_j) \text{ s.t. } x \in \{\pm 1\}^n$$

- **Solve the SDP relaxation:**

$$\max \frac{1}{2} \sum_{ij \in E} w_{ij}(1 - X_{ij}) \text{ s.t. } X \succeq 0, \ \mathrm{diag}(X) = e$$

  $\rightsquigarrow$ $v_1, \ldots, v_n$ unit vectors s.t. $X = (v_i^T v_j)$ is opt. for SDP.

- **Randomized rounding:** Pick a random hyperplane H with normal r.

  $\rightsquigarrow$ partition $(S, V \setminus S)$ depending on the sign of $v_i^T r$.

Performance analysis

Theorem: For $w \geq 0$,

$$\mathrm{mc}(G) \geq \underbrace{E(w(\delta(S))) \geq 0.878 \ \mathrm{sdp}(G)} \geq 0.878 \ \mathrm{mc}(G).$$

Basic lemma: $\mathrm{Prob}(ij \in \delta(S)) = \frac{\arccos(v_i^T v_j)}{\pi}$.

$$\begin{aligned} E(w(\delta(S))) &= \textstyle\sum_{ij \in E} w_{ij} \ \mathrm{Prob}(ij \in \delta(S)) \\ &= \textstyle\sum_{ij \in E} w_{ij} \ \frac{\arccos(v_i^T v_j)}{\pi} \\ &= \textstyle\sum_{ij \in E} w_{ij} \frac{(1 - v_i^T v_j)}{2} \ \underbrace{\frac{2}{\pi} \frac{\arccos(v_i^T v_j)}{1 - v_i^T v_j}}_{\geq \alpha_{\mathrm{GW}} \sim 0.878} \\ &\geq \mathrm{sdp}(G) \ \alpha_{\mathrm{GW}} \end{aligned}$$

Extension to ± 1 quadratic programming

Given $A \in \mathcal{S}_n$

Integer problem: $\mathrm{ip}(A) := \max \ x^T A x \ \text{ s.t. } \ x \in \{\pm 1\}^n$

SDP relaxation: $\mathrm{sdp}(A) := \max \ A \cdot X \ \text{ s.t. } \ X \succeq 0, \ \mathrm{diag}(X) = e.$

- $A = \frac{1}{4} L_w$, L_w: Laplacian matrix of (G, w) $\rightsquigarrow$ Max-Cut
  where $L_w(i,i) = w(\delta(i))$, $L_w(i,j) = -w_{ij}$.

$\rightsquigarrow$ When $A \succeq 0$, $Ae = e$, $A_{ij} \leq 0$ $(i \neq j)$ $\rightsquigarrow$ 0.878-approx. alg.

- When $A \succeq 0$ $\rightsquigarrow$ $\frac{2}{\pi}$ (~ 0.636)-approx. alg. [Nesterov 97]
- When $\mathrm{diag}(A) = 0$ $\rightsquigarrow$ Grothendieck constant

Nesterov $\frac{2}{\pi}$-approximation algorithm

- **Solve SDP:** Let $v_1, \ldots, v_n$ unit vectors s.t. $X = (v_i^T v_j)$ maximizes $A \cdot X$.
- **Random hyperplane rounding:** Pick a random unit vector r.

$\rightsquigarrow$ random ± 1 vector: $\quad x = (\mathrm{sgn}(r^T v_i))_{i=1}^n$

Lemma 1 [identity of Grothendieck] $E(xx^T) = \frac{2}{\pi} \arcsin X$.

Proof:

$$
\begin{aligned}
E(\mathrm{sgn}(r^T v_i)\ \mathrm{sgn}(r^T v_j)) &= 1 - 2\ \mathrm{Prob}(\mathrm{sgn}(r^T v_i) \neq \mathrm{sgn}(r^T v_j)) \\
&= 1 - 2\frac{\arccos(v_i^T v_j)}{\pi} = \frac{2}{\pi}\left(\frac{\pi}{2} - \arccos(v_i^T v_j)\right) \\
&= \frac{2}{\pi} \arcsin(v_i^T v_j).
\end{aligned}
$$

Global performance analysis

Lemma 2: $\arcsin X - X \succeq 0$

Proof: $\arcsin x - x = \sum_k a_k x^{2k+1}$ where $a_k \geq 0$.

Global analysis:

$$E(x^T A x) = A \cdot E(xx^T) = A \cdot (E(xx^T) - \frac{2}{\pi} X) + \frac{2}{\pi} A \cdot X$$

$$= \underbrace{\underbrace{A}_{\succeq 0} \cdot (\underbrace{\frac{2}{\pi} \arcsin X - \frac{2}{\pi} X}_{\succeq 0})}_{\geq 0} + \frac{2}{\pi} A \cdot X$$

$$\geq \frac{2}{\pi} A \cdot X.$$

Therefore: For $A \succeq 0$, $\mathrm{ip}(A) \geq \frac{2}{\pi}\, \mathrm{sdp}(A)$.

Grothendieck inequality

Assume $\mathrm{diag}(A) = 0$.
The **support graph** G_A has as edges the pairs ij with $A_{ij} \neq 0$.

Definition: The **Grothendieck constant** $K(G)$ of a graph G is the smallest constant K for which

$$\mathrm{sdp}(A) \;\leq K \;\; ip(A) \quad \text{for all} \;\; A \in \mathcal{S}_n \;\; \text{with } G_A \subseteq G.$$

Theorem: ([Gr. 53] [Krivine 77] [Alon-Makarychev(x2)-Naor 05])

- For G complete bipartite, $\frac{\pi}{2} \leq K(G) \leq \; \frac{\pi}{2}\frac{1}{\ln(1+\sqrt{2})} \sim 1.782$
- $\Omega(\log(\omega(G))) \leq K(G) \leq O(\log(\vartheta(\bar{G})))$.

Sketch of proof for Krivine's upper bound: $\frac{\pi}{2}\frac{1}{\ln(1+\sqrt{2})}$

Show: $K(K_{n,m}) \leq \frac{\pi}{2}\frac{1}{\ln(1+\sqrt{2})} =: \frac{\pi}{2c}$? $c := \ln(1+\sqrt{2})$

1. Let $A \in \mathbb{R}^{n\times m}$. Let u_i ($i \leq n$) and v_j ($j \leq m$) be unit vectors in H maximizing $\mathrm{sdp}(A) = \sum_{i\leq n,\ j\leq m} a_{ij} u_i \cdot v_j$.
2. **Construct** new unit vectors $S(u_i)$, $T(v_j) \in \hat{H}$ satisfying
$$\arcsin(S(u_i)\cdot T(v_j)) = c\ u_i \cdot v_j$$
3. Pick a random unit vector $r \in \hat{H}$. Define the ± 1 vectors x, y
$$x_i = \mathrm{sgn}(r^T S(u_i)),\quad y_j = \mathrm{sgn}(r^T T(v_j))$$
4. **Analysis:**
$$E(x^T A y) = \sum_{i,j} a_{ij}\ E(x_i y_j) = \sum_{i,j} a_{ij}\ \frac{2}{\pi}\arcsin(S(u_i)\cdot T(v_j))$$
$$= \sum_{i,j} a_{ij}\ \frac{2}{\pi}\ c\ u_i\cdot v_j = \frac{2}{\pi} c\ \mathrm{sdp}(A).$$

Proof (continued)

Step 2. Given unit vectors $u, v \in H$, construct $S(u), T(v) \in \hat{H}$ satisfying $\arcsin(S(u) \cdot T(v)) = c\ u \cdot v$.

- $\sin x = \sum_{k \geq 0} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$, $\sinh x = \sum_{k \geq 0} \frac{x^{2k+1}}{(2k+1)!}$
- Set $c := \sinh^{-1}(1) = \ln(1+\sqrt{2})$.
- Set $S(u) = \left(\sqrt{\frac{c^{2k+1}}{(2k+1)!}} u^{\otimes(2k+1)} \right)_k \in \hat{H} := \oplus_{k \geq 0} H^{\otimes(2k+1)}$,
  $T(v) = \left((-1)^k \sqrt{\frac{c^{2k+1}}{(2k+1)!}} v^{\otimes(2k+1)} \right)_k \in \hat{H}$.
- Then, $S(u) \cdot T(v) = \sum_k (-1)^k \frac{c^{2k+1}}{(2k+1)!} (u \cdot v)^{2k+1} = \sin(c\ u \cdot v)$.
  Thus: $\arcsin(S(u) \cdot T(v)) = c\ u \cdot v$.

A reformulation of the theta number

Theorem [Alon-Makarychev(x2)-Naor 05]
The smallest constant C for which

$$\mathrm{sdp}(-A) \leq \; C \;\mathrm{sdp}(A) \quad \text{for all } A \in \mathcal{S}_n \text{ with } G_A \subseteq G$$

is $C = \vartheta(\bar{G}) - 1$.

Geometrically:

- $\mathcal{E}_n := \{X \in \mathcal{S}_n \mid X \succeq 0, \; \mathrm{diag}(X) = e\}$ $\quad\rightsquigarrow$ the **elliptope**
- $\mathcal{E}(G) \subseteq \mathbb{R}^E$: the projection of $\mathcal{E}_n$ onto the edge set of G.
- **Theorem [AMMN]:** $-\mathcal{E}(G) \subseteq (\vartheta(\bar{G}) - 1) \; \mathcal{E}(G)$.

Link with matrix completion

Matrix completion: Given a partial $n \times n$ matrix, whose entries are specified on the diagonal (say **equal to 1**) and on a subset E of the positions (given by $x \in \mathbb{R}^E$), decide whether it can be completed to a PSD matrix.

Equivalently, decide whether $x \in \mathcal{E}(G)$?

- **A necessary condition:** Each fully specified principal submatrix is PSD. [Clique condition]
- The clique condition is sufficient IFF G is a chordal graph (i.e. no induced circuit of length ≥ 4).

$$\begin{pmatrix} 1 & 1 & a? & -1 \\ 1 & 1 & 1 & b? \\ a? & 1 & 1 & 1 \\ -1 & b? & 1 & 1 \end{pmatrix}$$ is not completable to PSD

Another necessary condition

Fact: $\begin{pmatrix} 1 & \cos a & \cos b \\ \cos a & 1 & \cos c \\ \cos b & \cos c & 1 \end{pmatrix} \succeq 0 \iff \begin{cases} a+b+c & \leq 2\pi \\ a-b-c & \leq 0 \\ -a+b-c & \leq 0 \\ -a-b+c & \leq 0 \end{cases}$

- Write $x = \cos a$ for some $a \in [0,\pi]^E$.
  If $x \in \mathcal{E}(G)$ then [Metric condition]

  $$a(F) - a(C \setminus F) \leq \pi(|F|-1) \;\; \forall\; C \text{ circuit}, \; F \subseteq C \text{ odd}.$$

- The metric condition is sufficient IFF G has no K_4 minor.

  $$\begin{pmatrix} 1 & -1/2 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 & -1/2 \\ -1/2 & -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & -1/2 & 1 \end{pmatrix} \not\succeq 0$$

  while $\frac{2\pi}{3}(1,1,1,1)$ satisfies the triangle inequalities.

Geometrically

- $\mathrm{CUT}^{\pm 1}(G) \subseteq \mathcal{E}(G)$, with equality IFF G has no K_3 minor.
- $\mathcal{E}(G) \subseteq \cos(\pi\ \mathrm{MET}^{01}(G))$, with equality IFF G has no K_4 minor.
- $\mathcal{E}(G) \subseteq \cos(\pi\ \mathrm{CUT}^{01}(G))$, with equality IFF G has no K_4 minor.
- The Goemans-Williamson randomized rounding argument shows: If $v_1, \ldots, v_n$ are unit vectors and $a_{ij} := \arccos(v_i^T v_j)$ are their pairwise angles, then

$$\sum_{1 \leq i < j \leq n} c_{ij} a_{ij} \leq \pi\ c_0$$

if $c \cdot z \leq c_0$ is any inequality valid for the cuts of K_n.

Extension to max k-cut [Frieze Jerrum 95]

Max k-cut: Given $G = (V, E)$, $w \in \mathbb{R}^E_+$, $k \geq 2$, find a partition $\mathcal{P} = (S_1, \ldots, S_k)$ maximizing $w(\mathcal{P}) = \sum_{e \in E | e \text{ is cut by } \mathcal{P}} w_e$.

- Pick unit vectors $a_1, \ldots, a_k \in \mathbb{R}^k$ with $a_i^T a_j = -\frac{1}{k-1}$ for $i \neq j$.

$\rightsquigarrow$ **Model max k-cut:**

$$\mathrm{mc}_k(G) = \quad \max \frac{k-1}{k} \sum_{ij \in E} w_{ij}(1 - x_i^T x_j) \quad \text{s.t.} \quad x_1, \ldots, x_n \in \{a_1, \ldots, a_k\}.$$

- **SDP relax.:** $$\mathrm{sdp}_k(G) = \max \frac{k-1}{k} \sum_{ij \in E} w_{ij}(1 - v_i^T v_j) \quad \text{s.t.} \quad v_i \text{ unit vectors}, \ v_i^T v_j \geq -\frac{1}{k-1}.$$

- **Randomized rounding:** Pick k independent random unit vectors $r_1, \ldots, r_k \rightsquigarrow$ partition $\mathcal{P} = (S_1, \ldots, S_k)$ where $S_h = \{i \mid v_i^T r_h \geq v_i^T r_{h'} \ \forall h'\}$.

Analysis

- The probability that edge ij is not cut, i.e., v_i, v_j are both closer to the same r_h, is equal to k times a function $f(v_i^T v_j)$.

$$E(w(\mathcal{P})) = \sum_{ij\in E} w_{ij}(1 - kf(v_i^T v_j))$$
$$= \sum_{ij\in E} w_{ij} \underbrace{\frac{1 - kf(v_i^T v_j)}{1 - v_i^T v_j}\frac{k}{k-1}}_{\geq \alpha_k := \min_{-\frac{1}{k-1}\leq t\leq 1} \frac{1-kf(t)}{1-t}\frac{k}{k-1}} \frac{k-1}{k}(1 - v_i^T v_j)$$
$$\geq \alpha_k \, \mathrm{sdp}_k(G).$$

- $\alpha_2 = \alpha_{GW} \sim 0.878$: GW approximation ratio for max-cut.
  $\alpha_3 = \frac{7}{12} + \frac{3}{4\pi^2} \arccos^2(-1/4) > 0.836$ [de Klerk et al.]
  $\alpha_{100} > 0.99$.