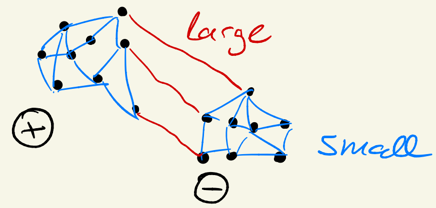
LECTURE 13



③ **Clustering** points $v_1, \dots, v_n \in \mathbb{R}^d$

$a_{ij} := \|v_i - v_j\|_2$

$$(*) \quad \boxed{\min_{x_i \in \{\pm 1\}} \sum_{i,j=1}^{n} a_{ij} x_i x_j} = \Sigma \text{distances in cluster } \oplus \text{ (where } x_i = 1) + \Sigma \text{distances in cluster } \ominus \text{ (where } x_i = -1) - 2\Sigma \text{distances between the clusters}$$

$\downarrow$

tries to make in-cluster distances small, across-cluster distances large.

- Works in $\forall$ metric space. [Frieze-Jerrum '1997]

SKIP ?

④ **Max cut** of graph $G = (V, E)$

$$\text{MaxCut}(G) = \max_{V = V^+ \cup V^-} |\overset{\text{cut}}{E(V^+, V^-)}| \;=$$

WLOG $V = \{1, \dots, n\}$; $A = [a_{ij}]$ is the adjacency matrix

$$a_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{if not} \end{cases} \qquad x_i = \begin{cases} 1 & \text{if } i \in V^+ \\ -1 & \text{if } i \in V^- \end{cases}$$

$$= \frac{1}{8} \max_{x_i \in \{\pm 1\}} \sum_{i,j=1}^{n} a_{ij} \underbrace{(x_i - x_j)^2}_{= \begin{cases} 4 & \text{if } (i,j) \in \text{same community} \\ 0, & \end{cases}}$$

(extra factor $\frac{1}{2}$ to avoid double counting of $(ij), (ji)$)

$x_i^2 - 2x_i x_j + x_j^2 = 2(1 - x_i x_j)$

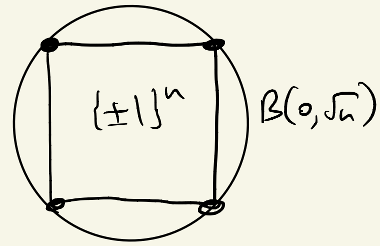
$$= \frac{1}{4} \max_{x_i \in \{\pm 1\}} \sum_{i,j=1}^{n} a_{ij} (1 - x_i x_j)$$

$$= \Sigma a_{ij} + \underbrace{\max_{x_i \in \{\pm 1\}} \Sigma (-a_{ij}) x_i x_j}_{= \, ?} \Rightarrow \text{equivalent to } (*)$$

SPECTRAL RELAXATION

$$\underset{\text{Boolean cube}}{\{-1,1\}^n} \subseteq \underset{\substack{\text{Euclidean ball}\\ \text{of radius } \sqrt{n}}}{B(0,\sqrt{n})}$$

Boolean cube $\subset$ Euclidean ball of radius $\sqrt{n}$



$$\max_{x\in\{\pm1\}^n} \sum_{ij} a_{ij} x_i x_j \le \max_{\|x\|_2 \le \sqrt{n}} \sum_{ij} a_{ij} x_i x_j = (\sqrt{n})^2 \max_{\|x\|_2\le 1} \underbrace{\sum_{ij} a_{ij} x_i x_j}_{= x^T A x} \qquad A=(a_{ij})$$

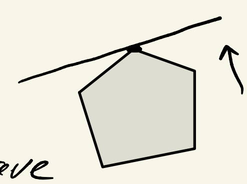
If A is symmetric,

$$= n\cdot \underbrace{\lambda_1(A)}_{\text{largest eigenvalue of } A.}$$

Maximizer x = the corresponding eigenvector of A. PCA.

- ☺ Sp.R is efficiently computable (diagonalize A)
- ☹ Tightness of SpR is unclear. No general guarantees.

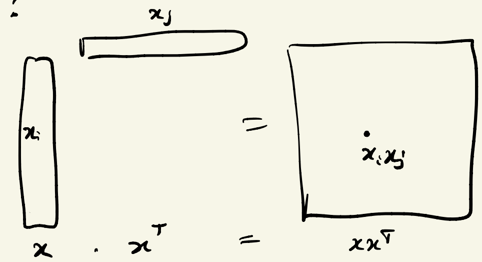
SEMIDEFINITE RELAXATION



- Convex program: maximize a linear function over a convex set.
  generally, concave
- Efficient solvers. Convex relaxation of

$$\max_{x_i\in\{\pm1\}} \sum_{ij} a_{ij} \boxed{x_i x_j} \;\; =: Z_{ij}$$

?



- Matrix $Z = [x_i x_j]_{i,j=1}^n = xx^T$
  - (a) symmetric
  - (b) $Z_{ii} = x_i^2 = 1 \quad \forall i$
  - (c) positive semidefinite $(Z \succeq 0)$: $u^T Z u \ge 0 \;\; \forall u\in\mathbb{R}^n$

$$\left[u^T x x^T u = \langle u,x\rangle\langle x,u\rangle = \langle x,u\rangle^2 \ge 0 \right]$$

- Hence, relax:

$$\max_{x_i \in \{\pm 1\}} \sum_{i,j} a_{ij} x_i x_j \le \boxed{\max_{\substack{Z \succeq 0, \\ Z_{ii}=1 \ \forall i}} \sum_{i,j=1}^{n} a_{ij} Z_{ij} = SDP(A)}$$

- The objective function $\sum_{ij} a_{ij} Z_{ij} = \langle A, Z \rangle$ is <u>linear</u> in Z
- The feasible set $\{ Z : \underbrace{Z \succeq 0}, \underbrace{Z_{ii} = 1 \ \forall i} \}$ is <u>convex</u>

(convex set "spectrahedron" $\cap$ linear subspace)

- $\Rightarrow$ SDP(A) is a convex program. "<u>Semidefinite program</u>"

Efficient solvers. (1000 variables are OK)

- <u>Geometric form of SDP</u>?

HW

- <u>Def/Prop</u> The Gram matrix of vectors $v_1, \dots, v_n \in \mathbb{R}^d$ is $[\langle v_i, v_j \rangle]_{i,j=1}^{n}$

It is always PSD.

Conversely, $\forall$ $n \times n$ PSD matrix is a Gram matrix of some vectors $v_1, \dots, v_n \in \mathbb{R}^n$.

$\Rightarrow$ in SDP(A), $\quad Z = [\langle v_i, v_j \rangle]$, $\quad Z_{ii} = \langle v_i, v_i \rangle = \|v_i\|_2^2 = 1 \ \forall i \quad \Rightarrow$

$$\boxed{SDP(A) = \max_{\substack{v_i \in \mathbb{R}^n \\ \text{unit}}} \sum_{ij} a_{ij} \langle v_i, v_j \rangle}$$

- Similarly, the SDP relaxation of

$$MaxCut(A) = \frac{1}{4} \max_{x_i \in \{\pm 1\}} \sum_{i,j=1}^{n} a_{ij} (1 - x_i x_j)$$

is

trivial $\wedge\!\!\!\!\wedge$? $\vee\!\!\!\!\vee$

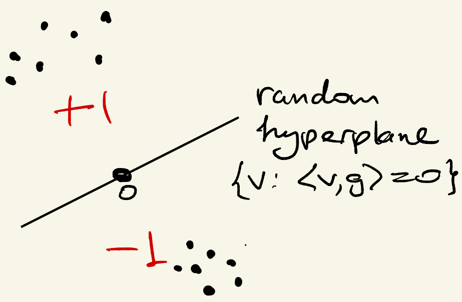
$$SDP(A) = \frac{1}{4} \max_{\substack{v_i \in \mathbb{R}^n \\ \text{unit}}} \sum_{i,j=1}^{n} a_{ij} \left(1 - \langle v_i, v_j \rangle\right).$$

And how can we convert solution v_i into a cut x_i ?

RANDOMIZED ROUNDING

- Let $g \sim N(0, I_n)$

$$v \mapsto \operatorname{sign}\langle v, g\rangle$$



- Does randomized rounding preserve geometry? (inner product)

Almost:

Lem (Grothendieck's identity) $\forall$ unit vectors $u, v \in \mathbb{R}^n$, $g \sim N(0, I_n)$

$$\mathbb{E}\, \operatorname{sign}\langle u, g\rangle \operatorname{sign}\langle v, g\rangle = \frac{2}{\pi} \arcsin\langle u, v\rangle.$$

HW

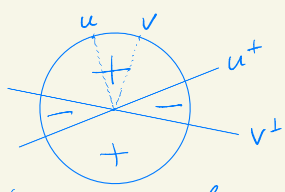
Rotation invariance

$\Rightarrow$ WLOG, $u, v \in \mathbb{R}^2$

$\Rightarrow$ $g \sim N(0, I_2)$

$g \mapsto g/\|g\|_2 =: \theta \sim \mathrm{Unif}(\text{circle})$. $\Rightarrow$ becomes a 2D problem.

Trigonometry $\Rightarrow$ QED



- ALGORITHM:

1. Solve $SDP(A) = \frac{1}{4} \max\limits_{v_i \in \mathbb{R}^n_{\text{unit}}} \sum\limits_{i,j=1}^{n} a_{ij}\left(1 - \langle v_i, v_j\rangle\right)$
2. Randomized rounding of the sol: $x_i := \operatorname{sign}\langle v_i, g\rangle$.
3. Output the cut $\{i : x_i = 1\} \mid \{j : x_j = -1\}$

Accuracy:

THM [Goemans-Williamson 1995]

The expected cut of this partition $\gtrsim 0.878 \cdot \mathrm{MaxCut}(G)$

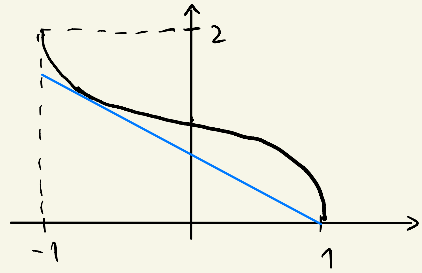
Proof Expected cut $= \mathbb{E}\, \frac{1}{4} \sum\limits_{i,j=1}^{n} a_{ij}\,(1 - x_i x_j)$

$$= \frac{1}{4} \sum_{i,j=1}^{n} a_{ij}\,(1 - \mathbb{E}\, x_i x_j)$$

$$1 - \mathbb{E}\, x_i x_j = 1 - \mathbb{E}\, \mathrm{sign}\langle v_i, g\rangle \langle v_j, g\rangle = 1 - \frac{2}{\pi}\arcsin\langle v_i, v_j\rangle$$

Construction

G.T.



$$= \frac{2}{\pi}\arccos\langle v_i, v_j\rangle \geq 0.878\,\big(1 - \langle v_i, v_j\rangle\big)$$

linearization: $\frac{2}{\pi}\arccos\theta \geq 0.878\,(1-\theta)$

$$\geq 0.878 \cdot \underbrace{\frac{1}{4}\sum_{ij} a_{ij}\big(1 - \langle v_i, v_j\rangle\big)}_{\parallel}$$

$$\mathrm{SDP}(A) \geq \mathrm{Max\,Cut}(A).$$

QED.

Remark

$$\mathrm{SDP}(A) = \frac{1}{4} \max_{\substack{v_i \in \mathbb{R}^n \\ \text{unit}}} \sum_{i,j=1}^{n} a_{ij}\big(1 - \langle v_i, v_j\rangle\big) = \frac{1}{8} \boxed{\max_{\substack{v_i \in \mathbb{R}^n \\ \text{unit}}} \sum_{i,j=1}^{n} a_{ij} \|v_i - v_j\|_2^2}$$

embeds the vertices of the graph into S^{n-1}.

connected vertices repel.