Sum-of-Squares and Max-Cut

Def (Max Cut): Given unweighted, undirected graph $G=(V,E)$, find $U \subseteq V$ maximizing $|E(U,\overline{U})|$.

Fact: Max-Cut is NP-hard.

Thm: there is a randomized $\frac{1}{2}$-approx for Max Cut.

Proof: Let $U \subseteq V$ be a random subset, i.e. $v \in U$ with prob $\frac{1}{2}$ $\forall v \in V$.
Each edge is cut (i.e. joins $E(U,\overline{U})$) with prob $\frac{1}{2}$.
So $|E(U,\overline{U})| \geq \frac{1}{2}|E| \geq \frac{1}{2}$ OPT.

LP Relaxation: $\max \sum_{ij \in E} d_{ij}$

s.t. $\{d_{ij} : i,j \in V\}$ form a metric

$d_{ij} \in [0,1] \quad \forall i,j \in V$

Fact: LP Relaxation has integrality gap 2. So cannot beat 2-approx.

Thm [Goemans-Williamson] 0.878-approx for Max-Cut by semi-definite programming.

This lecture: "modern" view of GW algo through Sum-of-Squares.

Consider the following polynomial optimization problem:

$$\max \sum_{ij \in E} (x_i - x_j)^2$$

$$\text{s.t. } x_i^2 = x_i \quad \forall i \in V$$

forces $x_i \in \{0,1\}$

$= 0$ if $x_i = x_j$

$= 1$ if $x_i \neq x_j$

So maximum is exactly OPT (no relaxation).

$= 1$ if $x_i \neq x_j$

So maximum is exactly OPT (no relaxation).

Cannot solve this program (NP-hard). How to relax?

<u>Sum-of-Squares (SoS) Relaxation:</u>

$$\max \ \tilde{\mathbb{E}}\Big[\sum_{ij \in E} (x_i - x_j)^2\Big]$$

s.t. $\tilde{\mathbb{E}}$ is a degree-d pseudo-expectation for Max-Cut

Def (Pseudo-expectation):

Let $\mathcal{P}_n^{\leq d}$ be the set of degree $\leq d$ polynomials over $x_1, \dots, x_n$.

A degree-d pseudo-expectation $\tilde{\mathbb{E}}$ is an operator $\tilde{\mathbb{E}} : \mathcal{P}_n^{\leq d} \to \mathbb{R}$ s.t.

1. $\tilde{\mathbb{E}}$ is linear: $\tilde{\mathbb{E}}[c_1 p_1 + c_2 p_2] = c_1 \tilde{\mathbb{E}}[p_1] + c_2 \tilde{\mathbb{E}}[p_2]$
2. $\tilde{\mathbb{E}}[1] = 1$
3. $\tilde{\mathbb{E}}[p^2] \geq 0$ for any $p \in \mathcal{P}_n^{\leq d/2}$ (to ensure $p^2 \in \mathcal{P}_n^{\leq d}$)

A degree-d pseudo-expectation for Max-Cut has the additional property

4. $\tilde{\mathbb{E}}[x_i^2 p] = \tilde{\mathbb{E}}[x_i p]$ for any $p \in \mathcal{P}_n^{d-2}$.

makes sense since $x_i^2 = x_i$ for ideal $x_i \in \{0,1\}$

Examples:

- For a fixed vector $y \in \mathbb{R}^n$, $\tilde{\mathbb{E}}[p] = p(y)$ is a pseudo-expectation.
- For vectors $y_1, \dots, y_m$ and scalars $a_1, \dots, a_m \geq 0$,
  $\tilde{\mathbb{E}}[p] = a_1 p(y_1) + a_2 p(y_2) + \cdots + a_m p(y_m)$ is a pseudo-expectation.

  Note that $\tilde{\mathbb{E}}[p^2] = a_1 p(y_1)^2 + \cdots + a_m p(y_m)^2$, hence "sum of squares"

Claim: SoS Relaxation is a relaxation.

Proof: Consider optimal $U \subseteq V$ and define $y \in \{0,1\}^n$: $y_i = 1$ if $i \in U$, $y_i = 0$ if $i \notin U$.

Define $\tilde{\mathbb{E}}$ as $\tilde{\mathbb{E}}[p] = p(y)$.

Define $\widetilde{\mathbb{E}}$ as $\widetilde{\mathbb{E}}[p] = p(y)$.
Then, objective value $\widetilde{\mathbb{E}}\left[\sum_{ij\in E}(x_i - x_j)^2\right] = \sum_{ij\in E}(y_i - y_j)^2 = OPT$.

Rounding an SoS Relaxation

Main Thm: There is a rounding algorithm that outputs (random) $U \subseteq V$ s.t.

$$\mathbb{E}\left[|E(U,\bar{U})|\right] \geq \alpha \cdot SoS \quad \text{where } \alpha \approx 0.878.$$

Claim: There exists optimal $\widetilde{\mathbb{E}}$ in SoS relaxation with $\widetilde{\mathbb{E}}[x_i] = 1/2 \ \forall i$.

Proof: Given $\widetilde{\mathbb{E}}$, construct $\widetilde{\mathbb{E}}'$ as

$$\widetilde{\mathbb{E}}'[p(x_1,\dots,x_n)] = \tfrac{1}{2}\widetilde{\mathbb{E}}[p(x_1,\dots,x_n)] + \tfrac{1}{2}\widetilde{\mathbb{E}}[p(1-x_1,\dots,1-x_n)],$$

So
$$\begin{aligned}\widetilde{\mathbb{E}}'[x_i] &= \tfrac{1}{2}\widetilde{\mathbb{E}}[x_i] + \tfrac{1}{2}\widetilde{\mathbb{E}}[1-x_i] \\ &= \tfrac{1}{2}\widetilde{\mathbb{E}}[x_i] + \tfrac{1}{2}\widetilde{\mathbb{E}}[1] - \tfrac{1}{2}\widetilde{\mathbb{E}}[x_i] \\ &= \tfrac{1}{2}.\end{aligned}$$

Define vector $\mu = \widetilde{\mathbb{E}}[x] = \begin{bmatrix}\widetilde{\mathbb{E}}[x_1] \\ \vdots \\ \widetilde{\mathbb{E}}[x_n]\end{bmatrix}$ and matrix $\Sigma = \widetilde{\mathbb{E}}[(x-\mu)(x-\mu)^T] = \begin{array}{c c} & j \\ i & \begin{bmatrix} & & \\ & \square & \\ & & \end{bmatrix}\end{array}$

$\widetilde{\mathbb{E}}[(x_i-\mu_i)(x_j-\mu_j)]$

Fact: For any positive semidefinite matrix Σ,
can sample n (dependent) Gaussians $y \in \mathbb{R}^n$ s.t.

$$\mathbb{E}[y] = \mu \text{ and } \underbrace{\mathbb{E}[(y-\mu)(y-\mu)^T]}_{\text{covariance matrix}} = \Sigma$$

Matrix Σ is positive semidefinite because for each $v \in \mathbb{R}^n$,

$$\begin{aligned} v^T\Sigma v = v^T\widetilde{\mathbb{E}}[(x-\mu)(x-\mu)^T]v &= \widetilde{\mathbb{E}}[v^T(x-\mu)(x-\mu)^T v] \\ &= \widetilde{\mathbb{E}}[\langle x-\mu, v\rangle^2] \geq 0.\end{aligned}$$

The algorithm samples y as above and performs threshold rounding:

The algorithm samples y as above and performs threshold rounding.

$U = \{ i \in V : y_i \leq \frac{1}{2} \}$.

We now analyze the approximation factor.

Define $\rho_{ij} = 4\,\tilde{\mathbb{E}}[x_i x_j] - 1 \quad \forall ij \in E$.

Claim: $\tilde{\mathbb{E}}[(x_i - x_j)^2] = \frac{1-\rho_{ij}}{2}$.

Proof:
$$\begin{aligned}
\tilde{\mathbb{E}}[(x_i-x_j)^2] &= \tilde{\mathbb{E}}[x_i^2] - 2\tilde{\mathbb{E}}[x_i x_j] + \tilde{\mathbb{E}}[x_j^2] \\
&= \tilde{\mathbb{E}}[x_i] - 2\tilde{\mathbb{E}}[x_i x_j] + \tilde{\mathbb{E}}[x_j] \\
&= \frac{1}{2} - \frac{1+\rho_{ij}}{2} + \frac{1}{2} \\
&= \frac{1-\rho_{ij}}{2}.
\end{aligned}$$

Claim: $\Pr[(i,j) \in E(U, \bar{U})] = \frac{\arccos(\rho_{ij})}{\pi}$.

Proof:
$$\begin{aligned}
\mathrm{Var}[y_i] &= \mathbb{E}[(x_i - \mu_i)^2] \\
&= \mathbb{E}[x_i^{\cancel{2}}] - 2\mu_i \mathbb{E}[x_i] + \mu_i^2 \\
&= \frac{1}{2} - 2\cdot\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{4} \\
&= \frac{1}{4}.
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}[y_i, y_j] &= \mathbb{E}[(x_i - \mu_i)(x_j - \mu_j)] \\
&= \mathbb{E}[x_i x_j] - \mu_i \mathbb{E}[x_j] - \mu_j \mathbb{E}[x_i] + \mathbb{E}[\mu_i \mu_j] \\
&= \frac{\rho_{ij} - 1}{4} - \frac{1}{2}\cdot\frac{1}{2} - \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} \\
&= \frac{\rho_{ij}}{4}.
\end{aligned}$$

Fact: given two Gaussians with covariance $\frac{\rho}{4}$,
$$\Pr[\mathrm{sign}(y_i - \mu_i) \neq \mathrm{sign}(y_j - \mu_j)] = \frac{\arccos\rho}{\pi}.$$

Thus, $\mathbb{E}[|E(u,\bar{u})|] = \sum_{ij\in E} \Pr[\text{sign}(y_i - \frac{1}{2}) \neq \text{sign}(y_j - \frac{1}{2})]$

$$= \sum_{ij\in E} \frac{\arccos\rho}{\pi}$$

$$= \sum_{ij\in E} \frac{\arccos\rho}{\pi} \cdot \frac{2}{1-\rho} \tilde{\mathbb{E}}[(x_i - x_j)^2]$$

$$\geq \min_{|\rho|\leq 1} \frac{2\arccos\rho}{(1-\rho)\pi} \cdot \text{SoS}.$$