

Sum of Squares: Part 2

Apoorv Vikram Singh

February 10, 2021

Recall MaxCut

- Given: $G = (V, E)$.
- Goal: Find $S \subseteq V$, such that $|E(S, \overline{S})|$ is maximized

Approximation Algorithm for MaxCut

- . Algorithm: Return a random cut.
- . In expectation: Algorithm cuts half the edges.
- . $\text{MaxCut} \leq |E|$.
- . Therefore, it is a $\frac{1}{2}$ -approximation algorithm.

Can we improve the $1/2$ -approximation?

- Question: Is there an LP-based algorithm that achieves $(0.5 + \varepsilon)$ -approximation algorithm?
- Answer: There does not exist a 2^{n^δ} size LP that gets $(0.5 + f(\delta))$ -approximation.
- . [Goemans-Williamson, 1994] Gave a 0.878-approximation algorithm for MaxCut (based on SDP).

Goal Today

- $G = (V, E)$, and let $\text{Opt}(G) = \text{MaxCut}(G)$.
- $f_G(\mathbf{x}) = \frac{1}{4} \sum_{(i,j) \in E} (\mathbf{x}_i - \mathbf{x}_j)^2$, for $\mathbf{x} \in \{-1, 1\}^n$.
- $\max_{\mathbf{x} \in \{-1, 1\}^n} f_G(\mathbf{x}) = \text{MaxCut}(G)$.

Theorem (0.878 Theorem)

For all G ,

$$\frac{\text{Opt}(G)}{0.878} - f_G(\mathbf{x}),$$

has a degree-2 SoS certificate.

To prove the theorem, we will prove a “rounding” theorem.

Theorem (Rounding Theorem)

Let μ be a degree-2 pseudo-distribution on $\{-1, 1\}^n$. Then, there is an actual distribution μ' such that

$$\mathbb{E}_{\mu'} f_G(\mathbf{x}) \geq 0.878 \tilde{\mathbb{E}}_{\mu} f_G(\mathbf{x}).$$

- Rounding: Takes pseudo-distribution to actual distribution.

Rounding Theorem \implies 0.878 Theorem

Proof.

Suppose $\frac{\text{Opt}(G)}{0.878} - f_G(\mathbf{x})$ is not SoS₂, then,

- \exists a degree-2 p.d. μ such that $\tilde{\mathbb{E}}_{\mu} \left(\frac{\text{Opt}(G)}{0.878} - f_G(\mathbf{x}) \right) < 0$.
- Rearranging: $\tilde{\mathbb{E}}_{\mu} f_G > \frac{\text{Opt}(G)}{0.878}$.
- Rounding Theorem $\implies \exists$ a distribution μ' , such that,

$$\mathbb{E}_{\mu'} f_G \geq 0.878 \tilde{\mathbb{E}}_{\mu} f_G(\mathbf{x}) > \text{Opt}(G).$$

- $\mathbb{E}_{\mu'} f_G > \text{Opt}(G)$, contradiction.



Interpreting Rounding Theorem

- Suppose we have a p.d. μ , and under this p.d.,
 $\mathbb{E}_{\mu} f_G(\mathbf{x}) = \text{Opt}_{\text{SoS}_2}$.
- We are interested in finding such cuts, or, if there are such cuts.
- Find distribution μ' , such that $\mathbb{E}_{\mu'} f_G(\mathbf{x})$ is as large as possible.
- We won't be able to prove it is equal, but we can prove

$$\mathbb{E}_{\mu'} f_G(\mathbf{x}) \geq 0.878 \text{Opt}_{\text{SoS}_2}.$$

- $\mu \rightarrow \mu'$ will be efficient \implies algorithm to approximate MaxCut.

Proving Rounding Theorem

Ideally:

- Given p.d. μ , find distribution μ' over $\{-1, 1\}^n$, such that

$$\mathbb{E}_{\mu'}(1, \mathbf{x})^{\otimes 2} = \tilde{\mathbb{E}}_{\mu}(1, \mathbf{x})^{\otimes 2}.$$

This is called: Generalized Moment Problem.

- Not possible**, otherwise we would have solved MaxCut exactly.

But, we can do it over \mathbb{R}^n

Lemma (Gaussian Sampling)

For any degree-2 p.d. μ , there exists an actual distribution over \mathbb{R}^n with same first and second moments.

Proof.

For any p.d. μ of degree-2,

$$\tilde{\mathbb{E}}_{\mu} \mathbf{x} \mathbf{x}^{\top} \succcurlyeq 0.$$

- First Moment: $\tilde{\mathbb{E}}_{\mu} \mathbf{x}$.
- Second Moment: $\tilde{\mathbb{E}}_{\mu} \mathbf{x} \mathbf{x}^{\top}$.
- Sample: $\mathbf{g} \sim \mathcal{N}(\tilde{\mathbb{E}}_{\mu} \mathbf{x}, \tilde{\mathbb{E}}_{\mu} \mathbf{x} \mathbf{x}^{\top})$.



Wlog $\tilde{\mathbb{E}}_{\mu} \mathbf{x} = \mathbf{0}$

- If μ was an actual distribution, then $\mathbf{x} \sim \mu$, and output $+\mathbf{x}$ or $-\mathbf{x}$ uniformly.
- Second Moment $\tilde{\mathbb{E}}_{\mu} \mathbf{x} \mathbf{x}^{\top}$ remains unchanged.
- Mean = $\mathbf{0}$.

Look at the p.d. with mean $\mathbf{0}$ and second moment $\tilde{\mathbb{E}}_{\mu} \mathbf{x} \mathbf{x}^{\top}$. The value of $\tilde{\mathbb{E}}_{\mu} f_G$ remains unchanged.

$$\begin{aligned} f_G(\mathbf{x}) &= \frac{1}{4} \sum_{(i,j) \in E} (\mathbf{x}_i - \mathbf{x}_j)^2 \\ &= \frac{1}{4} \sum_{(i,j) \in E} (2 - 2\mathbf{x}_i \mathbf{x}_j) \\ \implies \tilde{\mathbb{E}}_{\mu} f_G(\mathbf{x}) &= \frac{1}{4} \sum_{(i,j) \in E} (2 - 2\tilde{\mathbb{E}}_{\mu} \mathbf{x}_i \mathbf{x}_j). \end{aligned}$$

Efficient Algorithmic Process

- . $\mu \rightarrow \mathbf{g}$, such that $\tilde{\mathbb{E}}_{\mu} \mathbf{x} \mathbf{x}^{\top} = \mathbb{E} \mathbf{g} \mathbf{g}^{\top}$.
- . Issue: \mathbf{g} does not have entries in $\{\pm 1\}$.

Efficient Algorithmic Process,

1. Take $\mathbf{g} \sim \mathcal{N}(\mathbf{0}, \tilde{\mathbb{E}}_{\mu} \mathbf{x} \mathbf{x}^{\top})$.
2. $\hat{\mathbf{x}}_i = \text{sign}(\mathbf{g}_i)$, which gives that $\hat{\mathbf{x}} \in \{-1, 1\}^n$.

Call μ' the distribution on $\hat{\mathbf{x}}$.

Claim (Rounding Theorem)

$$\mathbb{E}_{\mu'} f_G(\mathbf{x}) \geq 0.878 \tilde{\mathbb{E}}_{\mu} f_G(\mathbf{x}).$$

Lemma (Sheppard's Lemma)

$$\mathbb{P}[\text{sign}(\mathbf{g}_i) \neq \text{sign}(\mathbf{g}_j)] \geq \frac{\arccos(\rho)}{2\pi(1-\rho)} \mathbb{E}(\mathbf{g}_i - \mathbf{g}_j)^2,$$

for $\rho = \tilde{\mathbb{E}}_{\mu} \mathbf{x}_i \mathbf{x}_j = \mathbb{E} \mathbf{g}_i \mathbf{g}_j$.

Remark(s): Comparing LHS and RHS of claim with lemma.

$$\cdot \mathbb{E}_{\mu'} f_G(\mathbf{x}) = \frac{1}{4} \sum_{(i,j) \in E} \mathbb{E}_{\mu'} (\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j)^2 = \frac{1}{4} \sum_{(i,j) \in E} \mathbb{P}[\text{sign}(\mathbf{g}_i) \neq \text{sign}(\mathbf{g}_j)] .$$

$$\cdot \tilde{\mathbb{E}}_{\mu} f_G(\mathbf{x}) = \frac{1}{4} \sum_{(i,j) \in E} \tilde{\mathbb{E}}_{\mu'} (\mathbf{x}_i - \mathbf{x}_j)^2 = \frac{1}{4} \sum_{(i,j) \in E} \mathbb{E}(\mathbf{g}_i - \mathbf{g}_j)^2 .$$

Sheppard's Lemma \implies Rounding Theorem

Proof.

$$\min_{\rho \in [-1,1]} \frac{\arccos(\rho)}{2\pi(1-\rho)} \geq \underbrace{\alpha_{GW}}_{=0.878...}, \quad (\text{min at } \rho = 0.69) .$$

This implies

$$\begin{aligned} \frac{1}{4} \mathbb{E}_{\mu'} (\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j)^2 &\geq \alpha_{GW} \frac{1}{4} \tilde{\mathbb{E}}_{\mu} (\mathbf{x}_i - \mathbf{x}_j)^2, \\ \frac{1}{4} \sum_{(i,j) \in E} \mathbb{E}_{\mu'} (\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j)^2 &\geq \alpha_{GW} \frac{1}{4} \sum_{(i,j) \in E} \tilde{\mathbb{E}}_{\mu} (\mathbf{x}_i - \mathbf{x}_j)^2. \end{aligned}$$



Proving Sheppard's Lemma

Proof

We have Gaussians $\mathbf{g}_i, \mathbf{g}_j$, such that $\mathbb{E} \mathbf{g}_i \mathbf{g}_j = \tilde{\mathbb{E}}_\mu \mathbf{x}_i \mathbf{x}_j = \rho$, and $\mathbb{E} \mathbf{g}_i^2 = \tilde{\mathbb{E}}_\mu \mathbf{x}_i^2 = 1$.

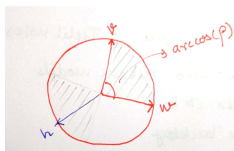
Procedure to generate such Gaussian vectors:

- . Let $\mathbf{v}, \mathbf{w} \in \mathbb{S}^{(2-1)}$ such that $\langle \mathbf{v}, \mathbf{w} \rangle = \rho$.
- . Take $\mathbf{h} \sim \mathcal{N}(\mathbf{0}, I_2)$.
- . $\hat{\mathbf{g}}_i = \langle \mathbf{h}, \mathbf{v} \rangle$, $\hat{\mathbf{g}}_j = \langle \mathbf{h}, \mathbf{w} \rangle$, this has same joint-distribution as $\mathbf{g}_i, \mathbf{g}_j$.

We are interested in

$$\mathbb{P}[\text{sign}(\mathbf{g}_i) \neq \text{sign}(\mathbf{g}_j)] = \mathbb{P}[\text{sign}(\hat{\mathbf{g}}_i) \neq \text{sign}(\hat{\mathbf{g}}_j)] .$$

Proof Cont...



$$\begin{aligned}\mathbb{P}[\text{sign}(\mathbf{g}_i) \neq \text{sign}(\mathbf{g}_j)] &= \mathbb{P}[\text{sign}(\hat{\mathbf{g}}_i) \neq \text{sign}(\hat{\mathbf{g}}_j)] \\ &= \mathbb{P}[\text{sign}(\langle \mathbf{h}, \mathbf{v} \rangle) \neq \text{sign}(\langle \mathbf{h}, \mathbf{w} \rangle)] \\ &= \frac{\arccos(\rho)}{\pi}.\end{aligned}$$

And the other quantity

$$\tilde{\mathbb{E}}_{\mu}(\mathbf{x}_i - \mathbf{x}_j)^2 = \mathbb{E}(\mathbf{g}_i - \mathbf{g}_j)^2 = \mathbb{E}(\hat{\mathbf{g}}_i - \hat{\mathbf{g}}_j)^2 = 2(1 - \rho).$$

$$\implies \mathbb{P}[\text{sign}(\mathbf{g}_i) \neq \text{sign}(\mathbf{g}_j)] \geq \frac{\arccos(\rho)}{2\pi(1 - \rho)} \mathbb{E}(\mathbf{g}_i - \mathbf{g}_j)^2$$



MaxCut Approximation Done.

Can we do better?

1. Can we do better with degree-2 SoS?: No.
2. Can we improve it with degree-4, degree-6, \dots , degree- $\log n$ SoS? Open.

How likely?

Unique Games Conjecture $\implies (\alpha_{GW} + \varepsilon)$ -approx to MaxCut is NP-Hard $\forall \varepsilon > 0$.

- Corollary: Suppose $\text{Opt}(G) \geq (1 - \delta) |E|$, then Gaussian rounding gives $\mathbb{E}_{\mu'} f_G(\mathbf{x}) \geq \left(1 - \mathcal{O}(\sqrt{\delta})\right) |E|$.
3. Is this the most optimal rounding? No (RPR² rounding does better in some regimes of δ).

Integrality Gaps?

What's the largest c for which degree-2 SoS certificate exists for $\frac{\text{Opt}(G)}{c} - f_G(\mathbf{x})$?

Ans: $c = 0.878..$ is optimal.

Fact

C_n : Cycle on n vertices, n odd.

$$\text{MaxCut}(C_n) = \text{Opt}(C_n) = \left(1 - \frac{1}{n}\right) |E|.$$

Theorem

There is a p.d. μ of degree-2 such that

$$\mathbb{E}_{\mu} f_{C_n}(\mathbf{x}) = \left(1 - \mathcal{O}\left(\frac{1}{n^2}\right)\right) |E|.$$

Choose $n = \frac{1}{\delta}$, then $\text{Opt}(C_n) = (1 - \delta) |E|$, and $\text{Opt}_{\text{SoS}_2}(C_n) \geq 1 - \mathcal{O}(\delta^2) |E|$.

\implies **Corollary** for small δ is tight up to constant factors.

Cycle = “Discretized” 2-dimn Sphere

\vdots

= “Discretized” high-dimn Sphere

[Feige-Schechtman'02] Proved α_{GW} is optimal.

Proof Sketch of Theorem

$$\text{MaxCut} = \max_{\mathbf{x} \in \{-1,1\}^n} \mathbf{x}^\top L_G \mathbf{x}.$$

$$\text{Related} = \max_{\|\mathbf{x}\|=\sqrt{n}} \mathbf{x}^\top L_G \mathbf{x} = n \|L_G\|_2.$$

- ▶ How to construct such a degree-2 p.d.?
 - Choose a distribution on \mathbf{x} that are in the “largest eigenspace” of L_G .
 - We just need $\tilde{\mathbb{E}}_\mu(1, \mathbf{x})(1, \mathbf{x})^\top \succcurlyeq 0$, $\tilde{\mathbb{E}}_\mu \mathbf{x}_i^2 = 1$, $\tilde{\mathbb{E}}_\mu \mathbf{1} = 1$.
- 1. Idea: $\lambda_{\max}(L_G) = 1 - \mathcal{O}(1/n^2)$. It is not Boolean because maxcut is $(1 - \mathcal{O}(1/n)) |E|$. Top eigenspace is 2-dimensional with vectors $\mathbf{v}_1, \mathbf{v}_2$.
- 2. set $M = \tilde{\mathbb{E}}_\mu \mathbf{x} \mathbf{x}^\top = \mathbf{v}_1 \mathbf{v}_1^\top + \mathbf{v}_2 \mathbf{v}_2^\top \succcurlyeq 0$.
- 3. Moreover, $\mathbf{v}_1 \mathbf{v}_1^\top + \mathbf{v}_2 \mathbf{v}_2^\top$ has diagonal entries 1.
- 4. Therefore, this is a valid pseudo-expectation.