CS369H: Hierarchies of Integer Programming Relaxations	Spring 2016-2017

Lecture 5. Max-cut, Expansion and Grothendieck's Inequality

Professor Moses Charikar — *Scribes: Kiran Shiragur*

Overview Here we derive SoS certificates for Maxcut, Expansion and Grothendieck's Inequality.

1 Recap

Lets restate some of the definitions and theorems needed for this class.

Definition 5.1. *(degree d pseudo-distribution) μ is a degree d pseudo-distribution iff*

- $\tilde{\mathbb{E}}_{\mu} 1 = 1$
- *For any polynomial f with degree $\leq d/2$* $\tilde{\mathbb{E}}_{\mu} f^2 \geq 0$

where $\tilde{\mathbb{E}}_{\mu} g := \sum_{x \in \{0,1\}^n} \mu(x) g(x)$

Lemma 5.1. *(polynomial representation of pseudo-distribution) For any degree d pseudo-distribution μ, there exists a multi-linear polynomial μ' of degree $\leq d$ such that:*

$$\tilde{\mathbb{E}}_{\mu} p = \tilde{\mathbb{E}}_{\mu'} p \quad \forall \text{polynomial } p \text{ of degree} \leq d$$

Lemma 5.2. *(moments) μ is a degree d pseudo-distribution iff*

- $\tilde{\mathbb{E}}_{\mu} 1 = 1$
- $\tilde{\mathbb{E}}_{\mu} \big((1, x)^{\otimes d/2}\big)\big((1, x)^{\otimes d/2}\big)^T \geq 0$

Lemma 5.3. *(duality between SoS proofs and pseudo-distribution) For any function $f : \{0,1\}^n \to \mathbb{R}$,*

$$\exists \text{ a degree } d \text{ SoS certificate for } f \iff \tilde{\mathbb{E}}_{\mu} f \geq 0 \text{ for every degree } d \text{ pseudo-distribution } \mu$$

One of the interesting facts about pseudo-distributions is that they behave closely to an actual distribution in the following sense. Most of the common interesting inequalities which are true for an actual distribution also hold for pseudo-distributions. For instance inequalities analogous to Cauchy-Schwarz and Holder's are also true for pseudo-distributions.

Lemma 5.4. *(Cauchy-Schwarz) For any degree d pseudo-distribution μ*

$$\left(\tilde{\mathbb{E}}_{\mu} PQ\right)^2 \leq \left(\tilde{\mathbb{E}}_{\mu} P^2\right)\left(\tilde{\mathbb{E}}_{\mu} Q^2\right)$$

where P and Q are polynomials of degree $\leq d/2$

Lemma 5.5. *For any degree 2 pseudo-distribution μ, there exist an actual distribution ρ over $\mathbb{R}^n$ such that:*

$$\tilde{\mathbb{E}}_{\mu}\left((1,x)^{\otimes 2}\right) = \mathbb{E}_{x\sim\rho}\left((1,x)^{\otimes 2}\right)$$

Proof. Let $\nu = \tilde{\mathbb{E}}_{\mu} x$ and $\Sigma = \tilde{\mathbb{E}}_{\mu}(x-\nu)(x-\nu)^T$. For any vector u:

$$\langle u, \Sigma u\rangle = \tilde{\mathbb{E}}_{\mu}\langle u, x-\nu\rangle^2 \geq 0 \quad \text{(Since } \mu \text{ is a degree 2 pseudo-distribution)} \Rightarrow \Sigma \succeq 0$$

Rest of the proof is pretty standard. Now pick a standard gaussian random vector g, and define a new random variable y to be:

$$y = \nu + \Sigma^{1/2} g$$

$$\mathbb{E}\, y = \nu$$

$$\mathbb{E}(y-\nu)(y-\nu)^T = \mathbb{E}\Sigma^{1/2} g g^T \Sigma^{1/2} = \Sigma^{1/2}(\mathbb{E} g g^T)\Sigma^{1/2} = \Sigma \qquad (\because \mathbb{E} g g^T = I)$$

□

In the rest of the lecture we derive SoS certificates for Max-cut, Expansion and Grothendieck's Inequality.

2 Max-cut

In this section we give SoS interpretation for Goemans-Williamson's 0.878 approximation algorithm for Max-cut. Lets first define the polynomial corresponding to max-cut:

$$f_G(x) = \sum_{(i,j)\in E(G)} (x_i - x_j)^2$$

Question 5.1. *How do we show an upper bound of c for $f_G(x)$?*

$$\max_{x\in\{0,1\}^n} f_G(x) \leq c$$

One way to show an upper bound is by showing the existence of SoS certificate for $c - f_G(x)$. In this lecture we show existence of a degree 2 SoS certificate for $c \geq \dfrac{\max f_G}{0.878}$, to be more precise we prove the following theorem.

Theorem 5.1. *For every graph $G(V,E)$, $\exists$ linear functions $g_1, g_2 \ldots g_n : \{0,1\}^n \to \mathbb{R}$ such that:*

$$\max(f_G) - 0.878 f_G(x) = \sum g_i(x)^2$$

To prove this result it suffices to prove the following equivalent theorem.

Theorem 5.2. *For every graph $G(V, E)$ and any degree 2 pseudo-distribution μ, there exists a probability distribution μ' over $\{0, 1\}^n$ such that:*

$$\mathbb{E}_{\mu'} f_G \geq 0.878 \tilde{\mathbb{E}}_\mu f_G$$

Proof. Without loss of generality $\tilde{\mathbb{E}}_\mu x = \frac{1}{2}\mathbf{1}$ (else consider the pseudo-distribution $\frac{1}{2}(\mu(x)+\mu(1-x))$). As earlier, construct a gaussian vector ξ with mean $\tilde{\mathbb{E}}_\mu x$ and co-variance $\tilde{\mathbb{E}}_\mu xx^T$. Lets calculate the variances first:

$$Var(x_i) = \tilde{\mathbb{E}}_\mu (x_i - 1/2)^2 = \tilde{\mathbb{E}}_\mu x_i^2 - 1/4 = 1/4 \quad \text{(Since } \mu \text{ is a pseudo-distribution over hypercube } \tilde{\mathbb{E}}_\mu x_i^2 = \tilde{\mathbb{E}}_\mu x_i)$$

$$Cov(x_i, x_j) = \tilde{\mathbb{E}}_\mu x_i x_j - 1/4$$

Define $\rho := 4Cov(x_i, x_j) = 4\tilde{\mathbb{E}}_\mu x_i x_j - 1$. With these definitions in mind, lets define x' as follows:

$$x_i' = 0 \quad \text{if } \xi_i < 1/2$$

$$x_i' = 1 \quad otherwise$$

Lets look at term wise expectation $\mathbb{E}(x_i' - x_j')^2$, and it is related to random variables $2\xi_i - 1$ and $2\xi_j - 1$ (converts variables into ± 1) in the following way.

$$\mathbb{E}(x_i' - x_j')^2 = \mathbb{P}\Big\{sign(2\xi_i - 1) \neq sign(2\xi_j - 1)\Big\} = \frac{2 \arccos \rho}{(1-\rho)\pi} \geq 0.878$$

□

One can show better results for instances where the Max-cut is close to all the edges.

Lemma 5.6. *For any degree 2 pseduo-distribution μ over $\{0, 1\}^2$ such that:*

$$\tilde{\mathbb{E}}_\mu \geq (1 - \epsilon)|E(G)|$$

there exists an actual distribution μ' over $\{0, 1\}^n$ with

$$\mathbb{E}_{\mu'} \geq (1 - 2\sqrt{\epsilon})|E(G)|$$

3 Expansion

For any d-regular graph $G(V, E)$, expansion of a set $S \subseteq V$ is defined as:

$$\phi_G(S) = \frac{|E(S, V \backslash S)|}{\frac{d}{n}|S||V \backslash S|}$$

Expansion ϕ_G of a graph $G(V, E)$ is:

$$\phi_G = \min_{S \subseteq V} \phi_G(S)$$

The goal here is to find a set $S \subseteq V$ which minimizes $\phi_G(S)$. If $x_i \in \{0,1\}$ denotes which side of the cut vertex i belongs and $f_G(x)$ defined as earlier then:

$$\phi_G = \min_{x \in \{0,1\}^n} \frac{f_G(x)}{\frac{d}{n}|x|(n-|x|)}$$

Note the function ϕ_G is of the form $\frac{P(x)}{Q(x)}$, where $P(x) = f_G(x)$ and $Q(x) = \frac{d}{n}|x|(n-|x|)$ are polynomials.

Question 5.2. *How do we certify* $\frac{P(x)}{Q(x)} \geq c$*?*

The above question is equivalent to certifying

$$P(x) - cQ(x) \geq 0$$

.

Lemma 5.7. $P(x) - \frac{1}{2}\phi_G^2 Q(x) \geq 0$ *has a degree* 2 *SoS certificate.*

The above lemma certifies the expansion of graph G is $\geq \frac{1}{2}\phi_G^2$.

Graph Expansion is a well studied problem and exhibits a rich literature.

[LR88] LP based algorithm which finds a set $S \subseteq V$ such that:

$$\phi_G(S) = O(\log n)\phi_G$$

[ARV09] SDP based algorithm (adds triangle inequalities to the standard SDP formulation) and achieves $O(\sqrt{\log n})$ approximation guarantee. It finds a set $S \subseteq V$ such that:

$$\phi_G(S) = O(\sqrt{\log n})\phi_G$$

This new SDP (with additional constraints) is captured by the degree 4 SoS.

One of the other famous results on expansion is the Cheeger's inequality.

Lemma 5.8. *(Cheeger's Inequality) For any* d*-regular graph* $G(V,E)$*, there exist a set* $S \subseteq V$ *and* $|S| \leq n/2$ *such that:*

$$\phi_G(S) \leq \sqrt{2\lambda}$$

where λ *is the second smallest eigenvalue of the normalized laplacian* $L_G = I - \frac{1}{d}A_G$*, and* A_G *is the adjacency matrix of graph* G*.*

Exercise 5.1. *Prove that Cheeger's inequality* $\Rightarrow$ *degree* 2 *SoS certificate for* $f_G(x) - \frac{1}{2}\phi^2(G) \cdot \frac{d}{n}|x|(n-|x|)$*.*

Hint: Note the quadratic form of Normalized laplacian can be represented in terms of $f_G(x)$.

$$\left\langle x, L_G x \right\rangle = \frac{1}{d} f_G(x)$$

4 Grothendieck's Inequality

Given a matrix $A \in \mathbb{R}^{n \times m}$, define

$$\|A\|_{\infty \to 1} := \max_{x \in \mathbb{R}^m} \frac{\|Ax\|_1}{\|x\|_\infty}$$

The above definition is equivalent to:

$$\|A\|_{\infty \to 1} = \max_{x \in \{\pm 1\}^m, y \in \{\pm 1\}^n} \Big\langle Ax, y \Big\rangle = \max_{x \in \{\pm 1\}^m} \|Ax\|_1$$

The quantity $\|A\|_{\infty \to 1}$ is important, for instance it approximates the cut-norm of a matrix within a constant factor:

$$cut - norm(A) = \max_{S \subseteq [n], T \subseteq [m]} \Big| \sum_{i \in S, j \in T} a_{ij} \Big| \in \Big[\frac{\|A\|_{\infty \to 1}}{4}, \|A\|_{\infty \to 1} \Big]$$

Theorem 5.3. *(Grothendieck's inequality) There exists a constant K_G such that for every matrix $A \in \mathbb{R}^{n \times m}$ and a degree 2 pseudo-distribution μ on $\{\pm 1\}^m \times \{\pm 1\}^n$.*

$$\tilde{\mathbb{E}}_{\mu(x,y)} \big\langle Ax, y \big\rangle \leq K_G \|A\|_{\infty \to 1}$$

In 1977, Krivine showed:

$$K_G \leq \frac{\pi}{2 \log(1 + \sqrt{2})} \approx 1.782$$

This result was later improved by [BMMN11]. Below we prove the Grothendieck's inequality due to Krivine.

Proof. Consider $\tilde{\mathbb{E}}_{\mu(x,y)} xy^T$, and suppose we could produce joint gaussian vectors ξ, ς such that:

$$\tilde{\mathbb{E}}_{\mu(x,y)} xy^T = K_{Krivine} \mathbb{E}[sign(\xi) sign(\varsigma)^T]$$

Then the quantity $\tilde{\mathbb{E}}_\mu \big\langle Ax, y \big\rangle$ is:

$$\tilde{\mathbb{E}}_\mu \big\langle Ax, y \big\rangle = Tr(\big\langle A, \tilde{\mathbb{E}}_\mu xy^T \big\rangle) = K_{Krivine} Tr(\big\langle A, \mathbb{E}[sign(\xi) sign(\varsigma)^T] \big\rangle) = K_{Krivine} \mathbb{E}[\big\langle A sign(\xi), sign(\varsigma) \big\rangle]$$

$$\leq K_{Krivine} \|A\|_{\infty \to 1}$$

Next we show existence of joint gaussian vectors ξ, ς with:

$$\tilde{\mathbb{E}}_{\mu(x,y)} xy^T = K_{Krivine} \mathbb{E}[sign(\xi) sign(\varsigma)^T]$$

For gaussian vectors:

$$\mathbb{E}[sign(\xi) sign(\varsigma)^T] = \frac{2}{\pi} \arcsin(\mathbb{E}[\xi \varsigma^T])$$ [1]

The goal here is to choose gaussian vectors ξ, ς such that:

$$\sin(c \tilde{\mathbb{E}}_{\mu(x,y)}[xy^T]) = \mathbb{E}[\xi \varsigma^T]$$

[1] Given a matrix A, the operations $\arcsin(A)$ and $\sin(A)$ apply functions arcsin and sin respectively entry wise

and this would give us $K_{Krivine} = \frac{\pi}{2c}$. Note the following matrix is PSD:

$$Cov = \begin{bmatrix} \left[\sinh(c\tilde{\mathbb{E}}xx^T)\right] & \left[\sin(c\tilde{\mathbb{E}}xy^T)\right] \\ \\ \left[\sin(c\tilde{\mathbb{E}}yx^T)\right] & \left[\sinh(c\tilde{\mathbb{E}}yy^T)\right] \end{bmatrix}$$

and we can choose gaussian vectors ξ, ς with Cov as the co-variance matrix, but we need the variances of entries of ξ, and ς to be 1. Note the matrices $\tilde{\mathbb{E}}xx^T$ and $\tilde{\mathbb{E}}yy^T$ have all the diagonal entries equal to 1 and the value of c should satisfy the equation $\sinh(c) = 1$, which evaluates to $c = \log(1 + \sqrt{2})$ which implies $K_{Krivine} = \frac{\pi}{2\log(1+\sqrt{2})}$. □

References

[ARV09] Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings and graph partitioning. *J. ACM*, 56(2):5:1–5:37, April 2009.

[BMMN11] M. Braverman, K. Makarychev, Y. Makarychev, and A. Naor. The grothendieck constant is strictly smaller than krivine's bound. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 453–462, Oct 2011.

[LR88] T. Leighton and S. Rao. An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms. In *[Proceedings 1988] 29th Annual Symposium on Foundations of Computer Science*, pages 422–431, Oct 1988.