Announcements

- start recording
- scribe: work w/other scribe to produce one set of notes
- My OH: Mon 12:30-1:30, Thu 9-10
- TF section/OH: Mon 2-3, Tue 8-9, Wed 5-6, Thu 5:30-6:30
- PS 4 posted. Project proposal due Sun 11/8, problems due TUE 11/17
- If Zoom goes down, check Piazza
- Sync whiteboard
- Post your project topic on spreadsheet
- Sign up for 15min check in with me this week or next
  → Fri 4:30-5:30, Mon 3-4, Tue 3-4, Wed 4-5, Thu 9

Agenda

- Recap
- Spectral Sparsification by Random Sampling
- Chernoff Bounds

RECAP

Loewner Order

- for symmetric $n \times n$ A, B: $A \succeq B$ iff $A - B$ is psd,
  i.e. $v^T A v \geq v^T B v$ for all $v \in \mathbb{R}^n$
- $A \succeq B \implies \lambda_k(A) \geq \lambda_k(B)$ for $k = 1, \dots, n$
- If A & B have common eigenbasis $v_1, \dots, v_n$ w/ e-values $\lambda_1, \dots, \lambda_n$ for A, e-values $\mu_1, \dots, \mu_n$ for B
  $\lambda_k \geq \mu_k$ for $k = 1, \dots, n \implies A \succeq B$

Expanders as Spectral Sparsifiers of Complete Graph

Prop: for a regular graph G w/ r.w. matrix W,

$$\gamma(G) \geq 1 - \omega \iff (1-\omega)(I-J) \preceq I - W \preceq (1+\omega)(I-J)$$

where J = all $1/n$ matrix

Cor: for every $\varepsilon > 0$, $\exists\ c = c(\varepsilon)$ such that for every n, there is a weighted graph H with $\leq cn$ edges (i.e. nonzero edge weights) such that

$$(1-\varepsilon) L(K_n) \preceq L(H) \preceq (1+\varepsilon) L(K_n)$$

Proof: Let G is a d_0-regular graph on n vertices w/ $\omega(G) \leq \frac{1}{2}$

Let $H = \frac{n}{d} G^{\log_2(1/\varepsilon)}$ $\qquad \omega(H) \leq \frac{1}{2^{\log_2(1/\varepsilon)}} = \varepsilon$

mult. all edge weights by $\frac{1}{d}$

H ~~d~~-regular $\underline{n}$ weighted degree, ~~for~~ $d = d_0^{\log_2(1/\varepsilon)} = \mathrm{poly}(\frac{1}{\varepsilon})$

has $\frac{d \cdot n}{2}$ edges

$c = d/2$.

Ramanujan graphs : $c = O(1/\varepsilon^2)$ $\left[\omega(G) \leq \frac{2\sqrt{d_0 - 1}}{d_0} = O\left(\frac{1}{\sqrt{d_0}}\right) \right]$

TODAY: Spectral Sparsifiers for Arbitrary Undirected Graphs

Thm: $\forall$ weighted undirected G on n vertices $\exists$ weighted undirected H w/ $O\left(\frac{n \log n}{\varepsilon^2}\right)$ edges

such that $(1-\varepsilon) L_G \preceq L_H \preceq (1+\varepsilon) L_G$

(Spielman Ch. 33: $O(n/\varepsilon^2)$ edges)

Stronger than a cut sparsifier (assuming "WLOG" $\vec{d}_H = \vec{d}_G$)

$$(1-\varepsilon)\phi_G(S) \leq \phi_H(S) \leq (1+\varepsilon)\phi_G(S)$$

$$\phi_H(S) = (x_S^T L_H x_S)/(x_S^T D_H x_S)$$

Slightly different notions of spectral approximation $H \approx_\varepsilon G$

		Exact Symmetry?	Exact Triangle Inequality
1)	$(1-\varepsilon) L_G \preceq L_H \preceq (1+\varepsilon) L_G$	$\times$	$\times$
2)	$(1+\varepsilon)^{-1} L_G \preceq L_H \preceq (1+\varepsilon) L_G$	$H \approx_\varepsilon G \Rightarrow G \approx_\varepsilon H$	$\times$
3)	$e^{-\varepsilon} L_G \preceq L_H \preceq e^{\varepsilon} L_G$	"	$H \approx_\varepsilon G,\ G \approx_\delta K \Rightarrow H \approx_{\varepsilon+\delta} K$

All equivalent up to a constant factor in ε for $\varepsilon \leq 1/2$

Proof of thm: Probabilistic Method

Include each edge e of G in H w.p. p_e, independently for each e

$$w_H(e) = \begin{cases} \frac{w_G(e)}{p_e} & \text{w.p. } p_e \\ 0 & \text{w.p. } 1-p_e \end{cases}$$

cf. take union of c random spanning trees ⇒ same marginal probs

$c = O\left(\frac{\log n}{\varepsilon^2}\right)$

$$p_e = c \cdot \ell_e = c \cdot w_G(e) \cdot R_{eff}(a,b) \quad \text{for } e=(a,b)$$

ℓ_e ← leverage score

$$R_{eff}(a,b) = (\delta_a - \delta_b)^T L^+ (\delta_a - \delta_b)$$

Assume for simplicity that $p_e \leq 1$ (else split e into multiple edges)

$$\mathbb{E}[\#\text{edges in } H] = \sum_e p_e = c \cdot \sum_e \ell_e = c \cdot (n-1)$$

$$\mathbb{E}[L_H] = \mathbb{E}\left[\sum_{e=(a,b)} w_H(e) \cdot L_{(a,b)}\right] \qquad L_{(a,b)} = (\delta_a - \delta_b)(\delta_a - \delta_b)^T$$

$$= \sum_{e=(a,b)} \mathbb{E}[w_H(e)] \cdot L_{(a,b)}$$

$$= \sum_{e=(a,b)} w_G(e) \cdot L_{(a,b)}$$

$$= L_G$$

<u>Goal</u>: show that whp over H

1) # edges in H $\leq 2cn$, and

2) $(1-\varepsilon) L_G \preceq L_H \preceq (1+\varepsilon) L_G$

$\Rightarrow \exists$ H satisfying (1) & (2).

<u>Chernoff Bound</u>: Let $X = \sum_{i=1}^{m} X_i$ for independent random variables X_i w/ $\Pr[X_i \in [0, R]] = 1$ and $\mu = \mathbb{E}[X]$. Then for all $0 < \varepsilon < 1$

$$\left.\begin{aligned} &\Pr[X \leq (1-\varepsilon)\mu] \leq e^{-\varepsilon^2 \mu / 2R} \\ \text{and } &\Pr[X \geq (1+\varepsilon)\mu] \leq e^{-\varepsilon^2 \mu / 3R} \end{aligned}\right\}$$

$\exp(-\Omega(m))$ for typical setting $\mu = \Omega(m)$, ε, R constant

"Sums of Bounded Independent RV's concentrate around expectation"

<u>Example</u>: $X = \#$ edges in $H = \sum_i X_e$ $X_e = \begin{cases} 1 & \text{if include edge } e \\ 0 & \text{o.w.} \end{cases}$

$\mu = \mathbb{E}[X] = \underline{c \cdot (n-1)}$ $X_e \in [0, 1]$

$R = 1$ $\varepsilon = 1/2$

$$\Pr[X \geq \underbrace{(1+\varepsilon)\mu}_{\frac{3}{2} c \cdot (n-1)}] \leq e^{-\varepsilon^2 \mu / 3R} = e^{-\Omega(cn)}$$

cf. Markov $\Pr[X \geq (1+\varepsilon)\mu] \leq \frac{1}{1+\varepsilon} = \frac{2}{3}$

Matrix Chernoff Bound $X = \sum_{i=1}^{m} X_i$, X_i independent random $n \times n$ psd matrices,
w/ $\Pr[\|X_i\| \le R] = 1$, $\mu_{min} = \lambda_{min}(E[X])$, $\mu_{max} = \lambda_{max}(E[X])$

$0 < \varepsilon < 1$

Then: $\Pr[\lambda_{min}(X) \le (1-\varepsilon)\mu_{min}] \le n \cdot e^{-\varepsilon^2 \mu_{min}/2R}$

$\Pr[\lambda_{max}(X) \ge (1+\varepsilon)\mu_{max}] \le n \cdot e^{-\varepsilon^2 \mu_{max}/3R}$

Application: $X = L_G^{+/2} L_H L_G^{+/2} = \sum_{e=(a,b)} X_e$

$$X_e = \begin{cases} \frac{w_e}{p_e} L_G^{+/2} L_{(a,b)} L_G^{+/2} & \text{w.p. } p_e \quad \text{(if we include edge } e\text{)} \\ 0 & \text{w.p. } 1-p_e \end{cases}$$

$$E[X] = E[L_G^{+/2} L_H L_G^{+/2}] = L_G^{+/2} E[L_H] L_G^{+/2}$$
$$= L_G^{+/2} L_G L_G^{+/2} = \Pi = \text{projection onto } \vec{1}^{\perp}$$

Restrict to $\vec{1}^{\perp}$: $\Pi \to$ identity, $\mu_{min} = \mu_{max} = 1$ assume wlog G is connected

wp 1

$$\|X_e\| \le \left\| \frac{w_e}{p_e} \cdot L_G^{+/2} L_{(a,b)} L_G^{+/2} \right\|$$
$$= \left\| \frac{w_e}{p_e} L_G^{+/2} (\delta_a - \delta_b)(\delta_a - \delta_b)^T L_G^{+/2} \right\| \qquad \|vv^T\| = v^T v$$
$$= \frac{w_e}{p_e} (\delta_a - \delta_b)^T L_G^{+/2} L_G^{+/2} (\delta_a - \delta_b)$$
$$= \frac{w_e}{p_e} R_{eff}(a,b)$$
$$= \frac{1}{c} \qquad p_e = c \cdot \ell_e$$

$$\frac{\mu_{max}}{R} = \mu_{min}/R = c$$

$\Rightarrow$ except with probability $\leq 2n\, e^{-\varepsilon^2 c/3} \leq \frac{1}{10}$, we have

$$\lambda_{\max}^{\vec{1}^\perp}\left(L_G^{+/2} L_H L_G^{+/2}\right) \leq 1+\varepsilon \qquad \text{setting } c = O\left(\frac{\log n}{\varepsilon^2}\right)$$

and $\lambda_{\min}^{\vec{1}^\perp}\left(L_G^{+/2} L_H L_G^{+/2}\right) \geq 1-\varepsilon$

$A \succcurlyeq B$

$C^T A C \succcurlyeq C^T B C$

Also have $\ker\left(L_G^{+/2} L_H L_G^{+/2}\right) = \operatorname{span}(\vec{1}) = \ker(\Pi)$

Thus $(1-\varepsilon)\Pi \preccurlyeq L_G^{+/2} L_H L_G^{+/2} \preccurlyeq (1+\varepsilon)\Pi$

$\Rightarrow$ $(1-\varepsilon) L_G^{1/2} \Pi L_G^{1/2} \preccurlyeq L_G^{1/2}\left(L_G^{+/2} L_H L_G^{+/2}\right) L_G^{1/2} \preccurlyeq (1+\varepsilon) L_G^{1/2} \Pi L_G^{1/2}$

$\Rightarrow$ $(1-\varepsilon) L_G \preccurlyeq L_H \preccurlyeq (1+\varepsilon) L_G$

Algorithmic Considerations

Apply L^+ in time T

$\Rightarrow$ compute $R_{eff}(a,b) = (\delta_a - \delta_b)^T L^+ (\delta_a - \delta_b)$ in time T

$\Rightarrow$ compute all R_{eff}'s in time $T \cdot m$

$\Rightarrow$ Sparsify in randomized time $T \cdot m$

Better: use Johnson-Lindenstrauss dimensionality reduction to compute all R_{eff}'s in time $T + O(m \cdot (\log n)/\varepsilon^2)$ [Spielman Ch. 14]