Sparsification

CS 292F

May 24, 2021

Lecture 14

Semidefinite order $\succeq$ (symmetric matrices)

- $A \succeq 0$ means $\lambda_i(A) \geq 0 \ \forall i$
- $A \preceq B$ means $B - A \succeq 0$

- $A \succeq 0 \iff x^T A x \geq 0 \ \forall x$
- $A \preceq B \iff x^T A x \leq x^T B x \ \forall x$
- $A \preceq B \Rightarrow \lambda_i(A) \leq \lambda_i(B) \ \forall i$ but not $\Leftarrow$

- $A \preceq B \Rightarrow CAC^T \preceq CBC^T$ for all C (not even square)

$\Leftarrow$ if C is invertible.

$$0 \preceq A \preceq B$$

"should $\Rightarrow$" $B^{-1}A \preceq I$ but not symmetric

but: $B^{1/2} = \sqrt{B}$ and $B^{+/2} = \sqrt{B^+}$

Then $B^{+/2} A B^{+/2}$ is symmetric ¿

same evals as $B^{-1}A$, so

- $0 \preceq A \preceq B \implies B^{+/2} A B^{+/2} \preceq I$

- $0 \preceq A \preceq B$ and same null space $\Rightarrow B^+ \preceq A^+$
(eg connected Laplacians)

Henceforth in this lecture, all matrices are connected Laplacians.

H is an ε-approximation of G

means $(1-\varepsilon)G \preceq H \preceq (1+\varepsilon)G$

A VERY POWERFUL STATEMENT.

① Eigenvalues are similar: for all i:

$$(1-\varepsilon)\lambda_i(G) \le \lambda_i(H) \le (1+\varepsilon)\lambda_i(G)$$

② Operation on a vector is similar:

$$\|G-H\| \le \varepsilon\|G\|$$

thus $\dfrac{\|Gx-Hx\|}{\|x\|} \le \|G-H\|$

and $\dfrac{\|Gx-Hx\|}{\|x\|\,\|G\|} \le \varepsilon$

③ Solutions of linear systems are similar:

$$\frac{1}{1+\varepsilon}G^{+} \preceq H^{+} \preceq \frac{1}{1-\varepsilon}G^{+} \Longrightarrow$$

$$\frac{\|G^{+}b - H^{+}b\|}{\|G^{+}\|\,\|b\|} \le \frac{\varepsilon}{1-\varepsilon}$$

(4) G and H are good preconditioners for each other

$$(1-\varepsilon) I \preceq H^{+/2} G H^{+/2} \preceq (1+\varepsilon) I$$

so the condition number of the preconditioned matrix:

$$\kappa_f = \frac{\lambda_n (H^{+/2} G H^{+/2})}{\lambda_2 (H^{+/2} G H^{+/2})} = 1 + \Theta(\varepsilon)$$

$\Rightarrow$ CG converges in $\Theta(1)$ iterations.

⑤ $R^{\text{eff}}_{(a,b)}$ is similar between all pairs of vertices:

(because $R^{\text{eff}}_{(a,b)_G} = (1_a - 1_b)^T G^+ (1_a - 1_b)$)

⑥ All cuts in the graph are similar.

If $A \subseteq V$ then $1_A^T G 1_A$

$= \sum_{ab \in E} c(a,b) \left(1_A(a) - 1_A(b)\right)^2$ = weight of cut between A and $V-A$.

and $1_A^T G 1_A$ is close to $1_A^T H 1_A$.

"cut sparsifier"

$\Leftarrow \not\Rightarrow$

"spectral sparsifier"

Expander $\approx K_n$

"Random graphs" are ε-approx of $\frac{d}{n} K_n$

INTUITION: both EXP & K_n have "no structure"

This intuition suggests that dence graphs with "structure" shouldn't have sparse approximat:

#graphs(n vtxs, m edges) : $\binom{\binom{n}{2}}{m}$ say $n < m < \frac{n^2}{4}$

$$\approx \frac{(n^2/2)^m}{m!} \sim n^{2m} = \underline{2^{2m \log n}}$$

Now : if $m \sim \frac{n^2}{4}$ $\quad \# = 2^{n^2 \log n / 2}$

if $m \sim n$ $\quad \# = 2^{2n \log n}$ → much less!

if $m \sim n \log n$ $\quad \# = 2^{\log^2 n}$ ↗

SO THIS SHOULD BE IMPOSSIBLE.

WRONG !!!

<u>All</u> graphs have ε-approximations with $O(n)$ edges.

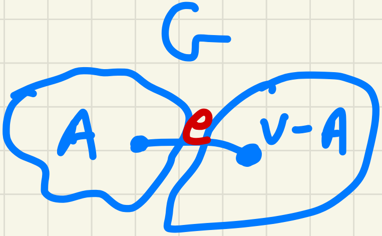
Ch 32: Gets $O(n \log n)$ edges with a very simple random sampling construction.

<u>DEF</u>: G a weighted Laplacian. weights $c(e) = c(a,b)$

The <u>leverage score</u> of edge $e = (a,b)$ is

$$l_e \overset{\text{DEF}}{\equiv} c(a,b) \cdot R^{\text{eff}}_{(a,b)} = \frac{R^{\text{eff}}_{(a,b)}}{r(a,b)}$$

$$= \frac{\text{effective resistance}}{\text{edge resistance}} \qquad 0 \le l_e \le 1$$



$$\ell_e = 1$$

leverage score

"importance" to resistive network.

<u>THM</u>: If we choose a spanning tree T with probability proportional to the product of its edge weights, the for every edge e

$$\Pr[e \in T] = \ell_e$$

Implies $\sum_{e \in E} \ell_e = n-1$

Sparsification: Given G, build $H \subseteq G$:

Pick edges at random with probability $p(a,b)$ of picking edge (a,b).

Then $c_H(a,b) = c_G(a,b)/p(a,b)$.

$$\mathbb{E}\, H = \sum_{a,b \in E(G)} p(a,b) \frac{c(a,b)}{p(a,b)} L_{(a,b)} = G$$

expectation (or mean)

$$G = \sum_{a,b \in E} c(a,b)\left(1_a - 1_b\right)\left(1_a - 1_b\right)^T$$

$$= \sum_{a,b \in E} c(a,b)\, L_{(a,b)}$$

① $p(a,b)$ proportional to l_{ab} to get $n \log n$ edges in H.

$$p(a,b) = \frac{\varepsilon^2}{3.5 \ln n}$$

KEY:

Matrix Chernoff bounds

showing whp

$$\lambda_i(H) \sim \lambda_i(\mathbb{E}(H)) = \lambda_i(G)$$