CS 598: Spectral Graph Theory. Lecture 18

Graph Sparsification by Effective Resistances

Alexandra Kolla

Today

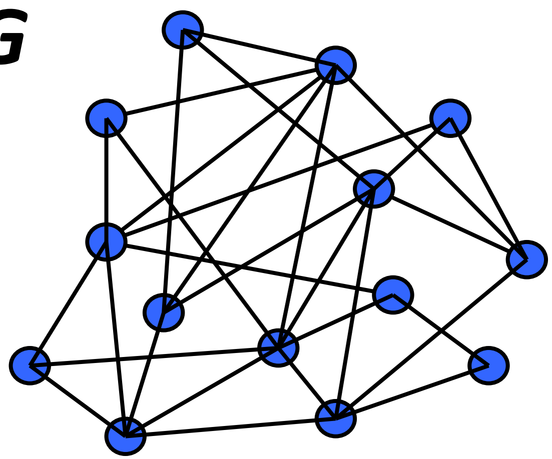
- Graph approximations.
- Sparsifiers for all graphs.
- Approximating a Projection Operator.
- Matrix Chernoff Bounds.

Sparsification

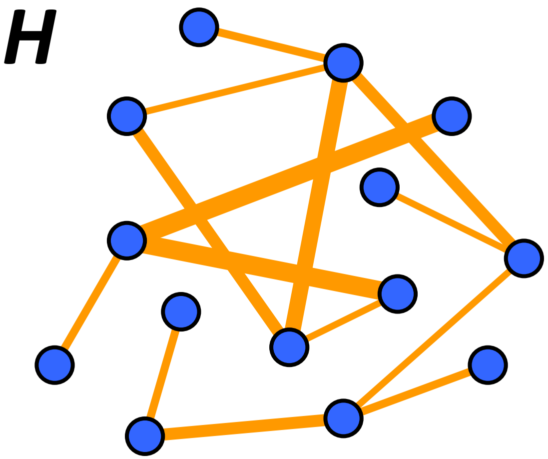
Approximate graph ***G*** by sparse graph ***H***

[BK'96],[ST'04],[SS'08],[BSS'09]

G



H



H is faster to compute with than ***G***

Can be used as proxy for G in computations

What kind of approximation?

Graphic Inequalities

- In Lecture 4, we defined

$$\mathrm{G} \succcurlyeq H \quad \text{if} \quad L_G \succcurlyeq L_H$$

Or, equivalently, $v^T L_G v \geq v^T L_H v$ for all v

- We say that G is an $\epsilon -$ approximation of H if

$$(1 - \epsilon)H \preccurlyeq \mathrm{G} \preccurlyeq (1 + \epsilon)H$$

Approximations of the Complete Graph

- Let G be a d-regular graph whose adjacency eigenvalues satisfy $|\alpha_i| \leq \epsilon d$.

- As its Laplacian eigenvalues satisfy $\lambda_i = d - \alpha_i$, all non-zero eigenvalues are between $(1 - \epsilon)$d and $(1 + \epsilon)d$.

- This means that for all x orthogonal to the all-one's vector

$$(1 - \epsilon)dx^T x \preccurlyeq x^T L_G x \preccurlyeq (1 + \epsilon)dx^T x$$

$$(1 - \epsilon)\frac{d}{n}K_n \preccurlyeq G \preccurlyeq (1 + \epsilon)\frac{d}{n}K_n$$

For $\epsilon \geq \frac{2}{\sqrt{d}}$

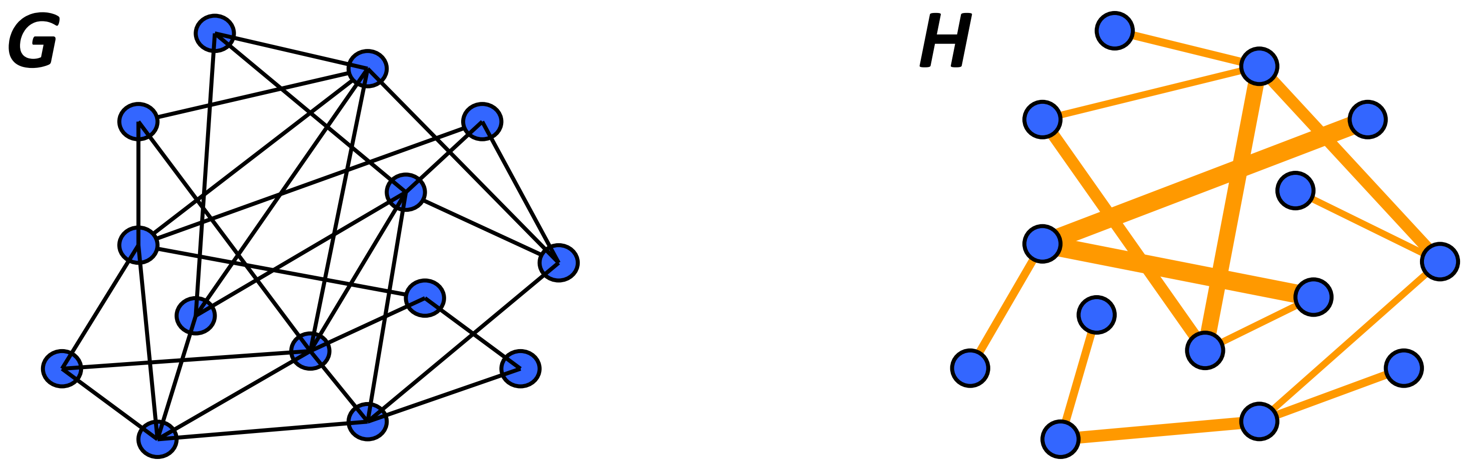
Sparsification

- We can approximate a graph with $O(n^2)$ edges (complete graph) by a graph with O(dn) edges (expander).

- Today: In general, we can ϵ −approximate any graph with another graph that has $O(\epsilon^{-2} n \log n)$ edges. (can be improved to $O(\epsilon^{-2} n)$ edges)

Sparsification

Approximate any graph ***G*** by sparse graph ***H!!!***



- ***H*** is faster to compute with than G
- Eigenvalues of H and G are similar
- Boundaries of all sets are similar (cuts)
- Effective Resistances are similar
- Solutions of linear equations in the two Laplacians are similar.

Motivation:Linear System Solving

- Want to solve $Ax = b$
- Assume $(1-\epsilon)A \preccurlyeq \mathrm{B} \preccurlyeq (1+\epsilon)A$
- B is easier to work with (sparser) than A
- $B^{-1}Ax = B^{-1}b \approx Ix = x$

Sparsification Plan

- Start with empty graph
- We will randomly sample edges:
- Create probability distribution over edges p_e, repeatedly use this to choose edges and add to the graph
- If we include an edge, add it with weight $1/p_e$.
- $\sum p_e = 1$
- Repeat q times, for some q tbd.

Sparsification Plan

- Equivalently to independently sampling matrices R_i, $\mathrm{i} = 1, \dots, \mathrm{q}$ such that $R_i = \frac{1}{\mathrm{p_e}} \mathrm{L_e}$ with probability $\mathrm{p_e}$.

- $E(R_i) = \sum_e 1\backslash \mathrm{p}_e\, p_e L_e = L = \mathrm{L_G}$
- Our sparsifier at the end is:

$$L_H = \frac{1}{q} \sum_{i=1 \ to\ q} R_i$$

Sparsification Plan

- $E(L_H) = \frac{1}{q}\sum_{i=1 \ to \ q} E(R_i) = L_G$

- We need to show that it is close with high probability. Would be true if we had real values random variables, but don't know what happens with matrix values R.V.

A Transformation

- H is an epsilon approximation of G iff

$$(1-\epsilon)x^T L_G x \preccurlyeq x^T L_H x \preccurlyeq (1+\epsilon)x^T L_G x$$

$$(1-\epsilon)G \preccurlyeq H \preccurlyeq (1+\epsilon)G \Rightarrow$$

$$-\epsilon\, G \preccurlyeq H - G \preccurlyeq \epsilon G \Rightarrow$$

$$-\epsilon\, \mathrm{L_G} \preccurlyeq L_G - L_H \preccurlyeq \epsilon L_G \Rightarrow$$

$$-\epsilon\, \Pi \preccurlyeq \Pi - L_G^{-1/2} L_H L_G^{-1/2} \preccurlyeq \epsilon \Pi \Rightarrow$$

$$\left\| L_G^{-1/2} L_H L_G^{-1/2} - \Pi \right\| \le \epsilon$$

A Transformation

$$\left\| L_G^{-1/2} L_H L_G^{-1/2} - \Pi \right\| \leq \epsilon$$

Where Π is projection orthogonal to the nullspace.

$$E\left(L_G^{-\frac{1}{2}} L_H L_G^{-\frac{1}{2}}\right) = \frac{1}{q} \sum_{i=1 \ to \ q} E(L_G^{-\frac{1}{2}} R_i L_G^{-\frac{1}{2}}) = L_G^{-\frac{1}{2}} L_G \, L_G^{-\frac{1}{2}} = \Pi$$

A Transformation

New plan is to sample matrices from

$$M_i = L_G^{-\frac{1}{2}} R_i L_G^{-\frac{1}{2}}$$

With some probability such that at the end, we get a matrix

$$M = \frac{1}{q} \sum \frac{1}{p_e} M_i$$

Which approximates the “identity”, in the sense that is it close w.h.p to E(M)= Π

Choosing the Probabilities

$$M = \frac{1}{q} \sum \frac{1}{p_e} M_i$$

Choose probabilities to be proportional to effective resistances.

$$p_e = \frac{1}{n-1} ||L_G^{-\frac{1}{2}} L_e L_G^{-\frac{1}{2}}|| = \frac{1}{n-1} b_e^T L_G^{-1} b_e$$
$$= \frac{1}{n-1} R_{eff}(e)$$

- Norm of each matrix $||M_i / p_e|| = n - 1$

Matrix Chernoff Bound

Theorem. Let Π be a projection matrix and M be a random psd matrix such that E(M)= Π and $\left|\left|M\right|\right| \leq \nu$. Let $M_1, \ldots, M_q$ be i.i.d from M. Then, for every $\epsilon > 0$

$$\Pr\left[||\frac{1}{q}\sum_i M_i - \Pi|| \geq \epsilon\right] \leq 2ne^{-\epsilon^2 q/4v}$$

Finish the result by taking $q = 5nl\mathrm{n}(2\mathrm{n})/\epsilon^2$ and observing that $v = n - 1$

Matrix Chernoff Bound

Theorem. Let Π be a projection matrix and M be a random psd matrix such that E(M)= Π and $||M|| \leq \nu$. Let $M_1, \dots, M_q$ be i.i.d from M. Then, for every $\epsilon > 0$

$$\Pr\left[||\frac{1}{q}\sum_i M_i - \Pi|| \geq \epsilon\right] \leq 2ne^{-\epsilon^2 q/4v}$$

Recall that for i.i.d mean-p bounded variables:

$\Pr[|\sum_i X_i - \mathrm{np}| \geq \epsilon n] \leq 2e^{-\epsilon^2 n/\sum(a_i - b_i)^2}$