LECTURES 1 & 2

- Understanding and computing electrical flows
- Approximating graphs and matrices via sampling

- Given an undirected graph G
  w. vertices V, $|V| = n$,
  edges E, $|E| = m$,
  edge weights $\underline{w} \in \mathbb{R}_+$,
- We pick arbitrary directions for each edge
- The edge-vertex incidence matrix is $B \in \mathbb{R}^{V \times E}$

  with $B(v,e) = -1$ for $e = (v,u)$ for some u
  $ +1$ for $e = (u,v)$ for some u
  $ 0$ o.w.

The Laplacian of G is $L = B \underbrace{W}_{\operatorname{diag}(\underline{w}) \in \mathbb{R}^{E \times E}} B^T$.

<u>Resistance</u> $r(e) = w(e)^{-1}$
$R = W^{-1}$

Flow conservation

flow $\underline{f} \in \mathbb{R}^E$

demand $\underline{d} \in \mathbb{R}^V$

$$B\underline{f} = \underline{d}$$ "the net flow at node $v \in V$ is $\underline{d}(v)$"

Electrical flow

Flow $$\underline{f}^* = \operatorname*{argmin}_{\substack{\underline{f} \in \mathbb{R}^E \\ B\underline{f} = \underline{d}}} \frac{1}{2}\underline{f}^T W^{-1} \underline{f}$$

Voltages $$\underline{f}^* = R^{-1}B^T \underline{x}^* = WB^T\underline{x}^*$$ (Ohm's law)

$$L\underline{x}^* = \underline{d},$$

$$\underline{x}^* = \operatorname*{argmin}_{\underline{x} \in \mathbb{R}^V} \frac{1}{2}\underline{x}^T L \underline{x} - \underline{d}^T \underline{x}$$

$$-\left(\frac{1}{2} x^{*T} L x^* - \underline{d}^T \underline{x}^*\right) = \frac{1}{2} f^* W^{-1} f^*$$ (Strong duality)

$\rightarrow$ Further reading: Lagrange duality
Exercise

Why solve $L\underline{x} = \underline{d}$?

In TCS & ML

- Random walks
- Graph structure from Laplacian eigenvalues & eigenvectors
  - → Cheeger's Inequality, graph decomposition
  - → Embedding graphs into low dimensional space.
    - ↓
  - → Semi-supervised learning
- Solving all kinds of flow problems
  - → Max flow, Min cost flow
  - → especially when making high accuracy algorithms
  - → CAVEAT: Maybe not in the future?
    - → Mixed 2, p norm flows?
- flow ideas ⟹ general convex optimization
  - → a common pipeline?
  - → max flow min cut duality → convex duality
  - → matrix approximation by row sampling
  - → accelerated convex optimization

Solving Laplacian Linear Equations

- Spielman & Teng 2004: nearly linear time
- tomorrow! → Kyng-Sachdeva 2016: very simple
- Jambulapati Sidford 2021: $O(|E| \text{poly} \log \log |E|)$
- → many other results: parallel, almost log space

Beyond TCS

- Scalar Elliptic Partial Differential Equations
  → Big topic in applied mathemathematics

Are Laplacian solvers practical?

- Heuristic solvers have been around for a long time
  - Incomplete Cholesky Factorization
  - (Algebraic) Multigrid
- Making TCS spectral graph theory useful in practice?
  - An open problem?!
  - Code? Laplacians.jl by Dan Spielman and others
  - A truly practical Laplacian solver?
    Ongoing effort ☺
  - Future: practical interior point via Laplacian solvers?

BEYOND FLOWS

fast solvers
- → Symmetric M-matrices
- → Directed Laplacians, general M-matrices
- → Connection Laplacians
- → Green's functions (Fast multipole methods)
- → 2D Truss problems (hard in general)

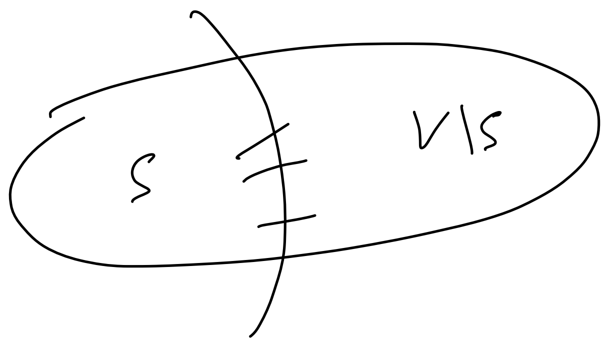
- Multi-commodity flow
  - → Linear equations as hard as general linear equations
  - → directed 2-commodity flow as hard as general linear programming

    LINK http://rasmuskyng.com/papers/LPto2CF.pdf
  - → many interesting questions in fine-grained complexity

Graph Approximations

- How to say that $G = (V, E_G, \underline{w}_G)$ is approximately equal to $H = (V, E_H, \underline{w}_H)$?
- Cut Inequality

Cut: $S \subseteq V$



Value

$$w(S) = \sum_{e \in E \cap (S \times V \setminus S)} w(e)$$

- Write $G \preceq_{cut} H$ if $\forall S \subseteq V.\ w_G(S) \le w_H(S)$.
- Write cG for the graph $(V, E_G, c\underline{w}_G)$

Theorem: Cut Sparsification (Benczur-Karger '96)

→ Given G, and $0 < \varepsilon < 1$, we can compute H s.t.

→ $(1-\varepsilon) G \leq_{cut} H \leq_{cut} (1+\varepsilon) G$

→ $|E_H| \leq O\left(\frac{n \log n}{\varepsilon^2} \right)$

$|V| = n$

$|E| = m$

→ The algorithm is randomized and succeeds with high probability and runs in time $O(m \log^2 n)$

→ NB: we assume $w_G(e) \in [1, poly(n)]$

Proof ideas

LEMMA (Karger '93)

- S'pose unweighted graph G has min cut c
- Then G has $\leq n^{2\alpha}$ cuts of size $\leq \alpha c$
- Find k connected components and sample
- Union bound

Spectral Sparsification

- $G \preceq H$ if $L_G \preceq L_H$ (Loewner Order)

LEMMA If $G \preceq H$ then $G \leq_{cut} H$.

Pf/ $\mathbb{1}_S \in \mathbb{R}^V$ indicator set.

Then $\mathbb{1}_S^T L_G \mathbb{1}_S = w_G(S)$

Spectral Sparsifier: H s.t. $\frac{1}{1+\varepsilon} G \preceq H \preceq (1+\varepsilon) G$

We abbreviate as $H \approx_\varepsilon G$

- Introduced by Spielman & Teng 2004
- Crucial for electrical flow algorithms (yields preconditioners)

Theorem Spectral Sparsification (Spielman & Srivastava 2008)

Given $G=(V,E,\underline{w})$ for any $0<\varepsilon<1$

and any $0<\delta<1$

there exist sampling probabilities $\underline{p}\in(0,1]^E$
such that if we for each edge e
we independently let

$$\begin{cases} e\in\tilde{E},\ \tilde{w}(e)=\frac{1}{\underline{p}(e)}w(e) & \text{w. prob } \underline{p}(e) \\ e\notin\tilde{E},\ \tilde{\underline{w}}(e)=\underline{0} & \text{o.w.} \end{cases}$$

then w. prob $\geq 1-\delta$,
we have for $\tilde{G}=(V,\tilde{E},\tilde{\underline{w}})$ that

$$\tilde{L}\approx_\varepsilon L \quad \text{(spectral graph approximation)}$$

and

$$|\tilde{E}|=O\left(\frac{n\log(n/\delta)}{\varepsilon^2}\right)$$

Theorem: Matrix Bernstein (Tropp 2010)

S'pose $X_1, \ldots, X_k \in \mathbb{R}^{n \times n}$ symmetric random matrices

zero-mean $\mathbb{E} X_i = 0_{n \times n}$

and independent

w. $\|X_i\| \le R$ $\qquad \|\cdot\|$: spectral norm

Let $X = \sum_i X_i$, $\qquad \sigma^2 = \|Var(X)\|$

$$= \|\sum_i \mathbb{E} X_i^2\|$$

Then

$$Pr\left[\|X\| \ge t\right] \le 2n \exp\left(\frac{-t^2}{2Rt + 4\sigma^2}\right)$$

$\boxed{A^2 \text{ means } A \cdot A, \text{ not entrywise!}}$

SANITY CHECK

Suppose each X_i has iid ± 1 entries on the diagonal and zero elsewhere.

$$X_i = n\left\{\overbrace{\begin{bmatrix} \pm 1 & & & \\ & \pm 1 & & \\ & & \ddots & \\ & & & \pm 1 \end{bmatrix}}^{n}\right.$$

diagonal

$$\|D\| = \max_i |D(i,i)|$$

$$\|X\| = \|\sum_i X_i\| = \max_j \left|\sum_i X_i(j,j)\right|$$

Theorem: Matrix Bernstein (Tropp 2010)

S'pose $X_1, \dots, X_k \in \mathbb{R}^{n \times n}$ symmetric random matrices

zero-mean $\mathbb{E} X_i = 0_{n \times n}$

and independent

w. $\|X_i\| \leq R$ $\qquad \|\cdot\|$: spectral norm

Let $X = \sum_i X_i$, $\qquad \sigma^2 = \|\operatorname{Var}(X)\|$
$= \|\sum_i \mathbb{E} X_i^2\|$

Then
$$\Pr\left[\|X\| \geq t\right] \leq 2n \exp\left(\frac{-t^2}{2Rt + 4\sigma^2}\right)$$

MARKOV'S LEMMA

S'pose $X \geq 0$ is a random variable.

Then $\Pr\left[X \geq t\right] \leq \frac{\mathbb{E} X}{t}$

aka $\Pr\left[X \geq c\,\mathbb{E} X\right] \leq \frac{1}{c} \qquad c > 1$

$\boxed{\exp(z) \text{ when } |z| \leq 1 \text{ ?}}$

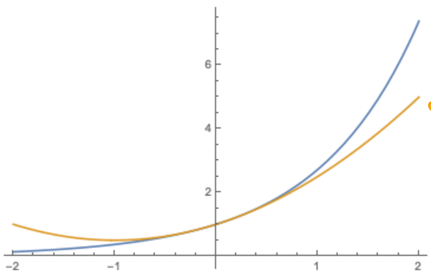
```
Series[Exp[z], {z, 0, 3}]
```

$1 + z + \frac{z^2}{2} + \frac{z^3}{6} + O[z]^4$

```
Plot[{Exp[z], 1 + z + z^2/2}, {z, -2, 2}]
```

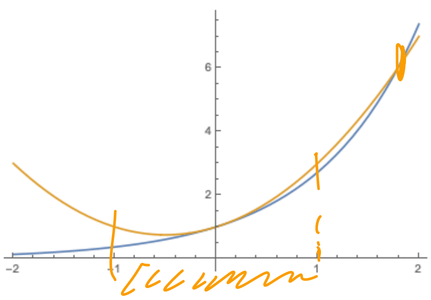
exp

taylor exp

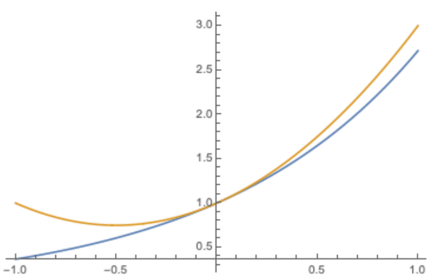


$\exp(z) \leq 1 + z + z^2$ for all $|z| \leq 1$

```
Plot[{Exp[z], 1 + z + z^2}, {z, -2, 2}]
```



```
Plot[{Exp[z], 1 + z + z^2}, {z, -1, 1}]
```



Proof of Matrix Bernstein ?

First we need some facts:

Trace: $tr(A) = \sum_i A(i,i) = \sum_i \lambda_i(A)$

Lemma If $f: \mathbb{R} \to \mathbb{R}$ is increasing then

$$A \to tr(f(A)) \text{ is increasing wrt. } \preceq$$

Lemma If $0 \prec A \preceq B$ then $\log(A) \preceq \log(B)$

$$\log(I + A) \preceq A \text{ for } A \succ -I$$

Lemma $\exp(A) \preceq I + A + A^2$ for symm. $A \in \mathbb{R}^{n \times n}$

$\exp(A) \succeq I + A$ w. $\|A\| \leq 1$

Lemma Jensen's Inequality

If $f: \mathbb{R} \to \mathbb{R}$ is convex, then $\mathbb{E} f(X) \geq f(\mathbb{E}X)$

.. is concave, then $\mathbb{E} f(X) \leq f(\mathbb{E}X)$

Theorem (Lieb)

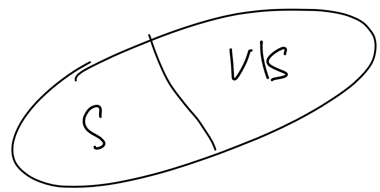
$$A \to tr(\exp(\underbrace{H}_{\text{symmetric real } H} + \log(A))) \text{ is concave over } A \succ 0.$$

A Very Different Spectral Sparsification Approach

Define $\mathrm{wdeg}(v) = \sum_{u \sim v} w(u,v)$

Conductance of a cut in G

$$\varphi_G(S) = \frac{w_G(S)}{\min\left(\mathrm{vol}_G(S),\ \mathrm{vol}_G(V \setminus S)\right)}$$



$$\varphi_G = \min_{S \subseteq V} \varphi_G(S) \longleftarrow \text{"Conductance of } G\text{"}$$

$$L = D - A$$

$$N = D^{-1/2} L D^{-1/2} = I - D^{-1/2} A D^{-1/2},$$

$$\lambda_1(N) = 0$$

Cheeger's Inequality

$$\frac{\lambda_2(N)}{2} \le \varphi(G) \le \sqrt{2\lambda_2(N)}$$

φ-Expander: a graph w. $\varphi_G \ge \varphi$

LEMMA: Cut to spectral approximation for expanders

- Let G be a φ-expander.
- Let H be a K-factor cut sparsifier of G:

$$\frac{G}{K} \preceq_{cut} H \preceq_{cut} KG$$

then H is a $\mathrm{poly}(K\varphi^{-1})$ spectral sparsifier of G.

Expander Decomposition

- Reducing general graph problems to problems on expanders
- Nearly linear time algo by Spielmang & Teng '04
- Cleaner & stronger result by Saranurak & Wang '19

Defn Induced subgraph: $G = (V, E)$, $U \subseteq V$

$$G[U] = (U, E \cap U \times U)$$

Theorem Given an unweighted graph $G=(V,E)$, $|E|=m$, there exists a randomised algorithm that whp. finds a partitioning of V into $V_1, \ldots, V_k$ s.t.

$$\forall i.\ \varphi_{G[V_i]} \geq \varphi$$

and $$\sum_i |E(V_i, V \setminus V_i)| = O(\varphi m \log^3 n).$$

Spectral Sparsification via Expander Decomp (ED) + Cut Sparsification

- Apply ED, repeat on graph w. edge set

$$E_{cut} = \bigcup_i |E(V_i, V \setminus V_i)|$$

$\rightarrow$ get decomp into $O\left(\frac{\log^4 n}{\varphi}\right)$ ☐? partitions into expanders

- Use Benczur-Karger cut sparsification on each expander
- Use return the union of cut sparsifiers

Proof Exercise.