Welcome back...

Metric spaces.

A metric space $X$, $d(i,j)$ where $d(i,j) \leq d(i,k) + d(k,j)$, $d(i,j) = d(j,i)$, and $d(i,j) \geq 0$.

Which are metric spaces?

(A) $X$ from $R^d$ and $d(\cdot, \cdot)$ is Euclidean distance.
(B) $X$ from $R^d$ and $d(\cdot, \cdot)$ is squared Euclidean distance.
(C) $X$-vertices in graph, $d(i,j)$ is shortest path distances in graph.
(D) $X$ is a set of vectors and $d(u,v)$ is $u \cdot v$.

Probabilistic Tree embedding.

Map $X$ into tree.

(i) No distance shrinks. (dominating)
(ii) Every distance stretches $\leq \alpha$ in expectation.

Map metric onto tree?

Algorithm

1. $\pi$ - random permutation of $X$.
2. Choose $\beta$ in $[\frac{1}{2}, 1]$
3. subtree$(S, \Delta)$:
   - $T = []$
   - if $\Delta < 1$ return $[S]$
   - foreach $i$ in $\pi$:
     - if $i \in S$:
       - $B = ball(i, \beta \Delta)$
       - $S = S \setminus B$
       - $T.append(B)$
       - return map $\lambda$: subtree$(x, \Delta/2)$, $T$
4. subtree$(X, D)$

Tree has internal node for each level of call. Tree edges have weight $\Delta$ to children.

Claim 1: $d_T(x,y) \geq d(x,y)$.

When $\Delta \leq d(x,y)$, $x$ and $y$ must be in different balls, so cut at lvl $\Delta \geq d(x,y)/2$.

$\Delta \geq d(x,y)$.

Approximate metric using a tree.

Tree metric:

$X$ is nodes of tree with edge weights $d_T(i,j)$ shortest path metric on tree.
Hierarchically well separated tree metric:
Tree weights are geometrically decreasing.

Probabilistic Tree embedding.

Map $X$ into tree.

(i) No distance shrinks. (dominating)
(ii) Every distance stretches $\leq \alpha$ in expectation.

Map metric onto tree?

Fix it up chappie!

For cycle, remove a random edge get a tree.

Stretch of edge: $\frac{n-1}{n} \times 1 + \frac{1}{2} \times (n-1) \approx 2$

General metrics?

Analysis: idea

Claim: $E[d_T(x,y)] = O(\log n)d(x,y)$.

Cut at level $\Delta \rightarrow d_T(x,y) \leq 4\Delta$. (Level of subtree call.)

$Pr[cut at level \Delta]$?

Would like it to be $\frac{d(x,y)}{n}$.

$\rightarrow$ expected length is $\sum_{\Delta}: 4\Delta \times \frac{d(x,y)}{n} = 4\log D \times d(x,y)$.

Why should it be $\frac{d(x,y)}{n}$?
Smaller the edge the less likely to be on edge of ball. Larger the delta, more room inside ball.

random diameter jiggles edge of ball.

$\rightarrow Pr[x \cdot y$ cut by $ball in ball] \approx \frac{d(x,y)}{D}$

The problem?
Could be cut be many different balls.

For each probability is good, but could be hit by many.
random permutation to deal with this...
The pipes are distinct!

\[ E(\delta(x, y)) = \Sigma_{\Delta \in D/2} \Sigma_{i \in X_\Delta} \left( \frac{1}{j} \right) 32d(x, y) \]

Recall \( X_\Delta \) has nodes with \( d(x, j) \in [\Delta/4, \Delta/2] \)

“Listen Stash, the pipes are distinct!!”

And now for something...

Completely different.
Edge Expansion/Conductance.

Graph G = (V, E),
Assume regular graph of degree d.

- **Edge Expansion.**
  \[ h(S) = \min_{(S, V - S)} \frac{|E(S, V - S)|}{|S||V - S|}, \quad h(G) = \min_S h(S) \]

- **Conductance.**
  \[ \phi(S) = \frac{|E(S, V - S)|}{|S||V - S|}, \quad \phi(G) = \min_S \phi(S) \]

Note \( n \geq \max(|S|, |V - S|) \geq n/2 \)
\[ h(G) \leq \phi(G) \leq 2h(S) \]

Spectra of the graph.

- **Eigenvector:** \( v - Mv = \lambda v \)  
  - Real, symmetric.
  - **Claim:** Any two eigenvectors with different eigenvalues are orthogonal.
  - **Proof:** Eigenvectors: \( v, v' \) with eigenvalues \( \lambda, \lambda' \).
    \[ v^T M v' = v^T (\lambda v') = \lambda' v^T v' \]
    \[ v^T M v = \lambda v^T v = \lambda' v^T v \]
  - Distinct eigenvalues \( \rightarrow \) orthonormal basis.

In basis: matrix is diagonal.
\[ M = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \]

Rayleigh Quotient

\[ \lambda_2 = \max_{x \neq 0} \frac{x^T M x}{x^T x} \]
- In basis, \( M \) is diagonal.

- Represent \( x \) in basis, i.e., \( x_i = x \cdot v_i \).
- \[ x^T M x = \sum_{i,j} x_i \lambda_i x_j \leq \lambda_1 x^T x = \lambda_1 \]
- Tight when \( x \) is first eigenvector.
- **Rayleigh quotient.**
  \[ \lambda_2 = \max_{x \neq 0 \perp 1} \frac{x^T M x}{x^T x} \]
  \[ x \perp 1 \Rightarrow x_i = 0. \]
  Example: \( 0/1 \) indicator vector for balanced cut, \( S \) is one such vector.
- Rayleigh quotient is \( \frac{|E(S, V - S)|}{2|S|(|V - S|)} \)
  \[ h(G) = \min_S \frac{|E(S, V - S)|}{2|S|(|V - S|)} \]

- **Cheeger’s inequality.**
  \[ \frac{1}{2} \leq h(G) \leq \sqrt{2(1 - \lambda_2)} \]
  Hmmm...

Connected \( \lambda_2 < \lambda_1 \).
\[ h(G) \text{ large} \rightarrow \text{well connected} \rightarrow \lambda_1 - \lambda_2 \text{ big.} \]
Disconnected \( \lambda_2 > \lambda_1. \)
Easy side of Cheeger.

Small cut $\rightarrow$ small eigenvalue gap.

$\frac{\mu_2}{2} \leq h(G)$

Cut $S$. $i \in S : v_i = |V| - |S|, i \notin S : v_i = |S|.$

$\sum_i v_i = |S||(V|-|S|)-|S||(V|-|S|) = 0$

$\rightarrow v \perp 1.$

$v^T v = |S||V|-|S| + |S|^2 |V|-|S| = |S||V|-|S|)(|V|)$.  

$v^T Mv = \frac{1}{2} \sum_{i \in V} x_i x_i.$

Same side endpoints: like $v^T v$.

Different side endpoints: $-|S||V|-|S|)$

$v^T Mv = v^T v - (2E(S, S)||S|(|V|-|S|))$

$\lambda_2 \geq 1 - 2h(S) \rightarrow h(G) \geq \frac{1-\lambda_2}{\lambda_2}$

See you ...

Thursday.