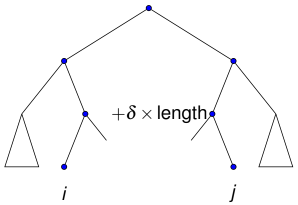
Tree

Path: $G = (V, E)$, $V = \{1, \ldots, n\}$, $E = \{(i, i+1) : i \in \{1, n-1\}\}$

Init: $f(e) = 0$, $\forall e \in G$.

Update(i, j, δ)
$\forall k \in \{i, \ldots, j-1\}$, $f(i, i+1) = f(i, i+1) + \delta$

Lookup(i, j)
Return $\sum_{k=i}^{j-1} f(k, k+1)$.



Leaves are edges.
Update: find common ancestor.
For left path:
Add $\delta \times$ length to edge.
Reflect for right path.

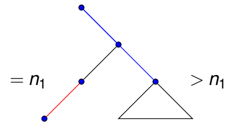
Lookup: Exercise.

$O(\log n)$ update/lookup.

Path Decomposition of Tree

Given tree T.

Decomposition Procedure:
Longest path from root to leaf.
From root, go towards heavier side.
Remove. Recurse



Decomposition Property:
Where a root to leaf path in tree only sees $O(\log n)$ paths.

Proof Idea:
Every path change, doubles number of vertices.

$O(\log^2 n)$ update time for updating flow values on tree edges!

What's going on?

Geometric View.
Cycles are constraints.
Flow around cycle $= 0$.
Each cycle update is approximate projection
into subspace defined by constraint.
Solution is intersection of cycle constraints and flow conservation.
Kirchhoff's Laws.

Algorithmic Power.
Algebra: Solving exactly on tree.
Calculus: make a local move th decreases potential.

Better Algorithm:
Recursive algorithm give $O(m\sqrt{\log n})$ iterations to halve error.
Preconditioner uses tree to make sparse version of graph.

Correspondence to Practice:
Random sparsification of Cholesky factorization.
(Kyung-Sachdeva)
Tree pre-conditioner.

Laplacian Systems are quite general:
Climate, physics, SDD-matrices.

Next

Rayleigh quotient.
$\lambda_2 = \max_{x \perp 1} \frac{x^T M x}{x^T x}$.

Eigenvalue gap: $\mu = \lambda_1 - \lambda_2 = \frac{1}{d}\lambda_{min}(L)$.

Recall: $h(G) = \min_{S, |S| \leq |V|/2} \frac{|E(S, V-S)|}{|S|}$

$\frac{\mu}{2} = \frac{1-\lambda_2}{2} \leq h(G) \leq \sqrt{2(1-\lambda_2)} = \sqrt{2\mu}$
Today: the right hand inequality.

Sum of squared differences.

Quotient Rayleigh: $\max_{x \perp 1} \frac{x^T M x}{x^t x}$.

Alternatively: $\mu = \min_{x \perp 1} \frac{x^T (L/d) x}{x^t x}$.

$L = dI - A$ or $L/d = I - A/d = I - M$.

Also:
$x^T L x = \sum_i d x_i^2 - \sum_{e=(i,j)} 2 x_i x_j$
$= \sum_{e=(i,j)} (x_i^2 + x_j^2 - 2 x_i x_j)$
$= \sum_{e=(i,j)} (x_i - x_j)^2$.

Denominator:
$\sum_i |x_i|^2 = |x|$. Normalize to 1.

Also note:
$\sum_{i,j} (x_i - x_j)^2 = \sum_{i,j} x_i^2 + x_j^2 - 2 x_i x_j = 2n|x|^2$.

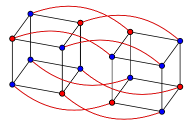
Since $x_j \sum_i x_i = 0$ for $x \perp 1$.

Rayleigh Quotient:
embedding to line
that minimizes ratio of Average edge length/(avg pair length)

Hypercube

$V = \{0, 1\}^d$ $(x, y) \in E$ when x and y differ in one bit.

$|V| = 2^d$ $|E| = d 2^{d-1}$.



Good cuts? "Coordinate cut": d of them.
Edge expansion: $\frac{2^{d-1}}{d 2^{d-1}} = \frac{1}{d}$

Ball cut: All nodes within $d/2$ of node, say $00 \cdots 0$.
Vertex cut size: $\binom{d}{d/2}$ bit strings with $d/2$ 1's.
$\approx \frac{2^d}{\sqrt{d}}$
Vertex expansion: $\approx \frac{1}{\sqrt{d}}$.
Edge expansion: $d/2$ edges to next level. $\approx \frac{1}{2\sqrt{d}}$

Worse by a factor of $\sqrt{d}$

Eigenvalues of hypercube.

Anyone see any symmetry?

Coordinate cuts. +1 on one side, -1 on other.

$(Mv)_i = (1-2/d)v_i$.

Eigenvalue $1-2/d$. d Eigenvectors. Why orthogonal?

Next eigenvectors?

Delete edges in two dimensions.

Four subcubes: bipartite. Color ± 1

Eigenvalue: $1-4/d$. $\binom{d}{2}$ eigenvectors.

Eigenvalues: $1-2k/d$. $\binom{d}{k}$ eigenvectors.

Back to Cheeger.

Coordinate Cuts:
Eigenvalue $1-2/d$. d Eigenvectors.

$\frac{\mu}{2} = \frac{1-\lambda_2}{2} \le h(G) \le \sqrt{2(1-\lambda_2)} = \sqrt{2\mu}$

For hypercube: $h(G) = \frac{1}{d}$ $\lambda_1 - \lambda_2 = 2/d$.
Left hand side is tight.

Note: hamming weight vector also in first eigenspace.

Lose "names" in hypercube, find coordinate cut?

Find coordinate cut?

Eigenvector v maps to line.
Cut along line.

Eigenvector algorithm gets a linear combination of coordinate cuts.

Something like ball cut.

Find coordinate cut?

Cycle

Tight example for Other side of Cheeger?

$\frac{\mu}{2} = \frac{1-\lambda_2}{2} \le h(G) \le \sqrt{2(1-\lambda_2)} = \sqrt{2\mu}$

Cycle on n nodes.

Will show other side of Cheeger is tight.

Edge expansion:Cut in half.
$|S| = n/2$, $|E(S,\overline{S})| = 2$
$\rightarrow h(G) = \frac{2}{n}$.

Show eigenvalue gap $\mu \le \frac{1}{n^2}$.

Find $x \perp \mathbf{1}$ with Rayleigh quotient, $\frac{x^T Mx}{x^T x}$ close to 1.

Slow vector.

Find $x \perp \mathbf{1}$ with Rayleigh quotient, $\frac{x^T Mx}{x^T x}$ close to 1.

$$x_i = \begin{cases} i - n/4 & \text{if } i \le n/2 \\ 3n/4 - i & \text{if } i > n/2 \end{cases}$$

Hit with M.

$$(Mx)_i = \begin{cases} -n/4 + 1/2 & \text{if } i = 1, n \\ n/4 - 1 & \text{if } i = n/2 \\ x_i & \text{otherwise} \end{cases}$$

$\rightarrow x^T Mx = x^T x(1 - O(\frac{1}{n^2})) \quad \rightarrow \lambda_2 \ge 1 - O(\frac{1}{n^2})$

$\mu = \lambda_1 - \lambda_2 = O(\frac{1}{n^2})$

$h(G) = \frac{2}{n} = \Theta(\sqrt{\mu})$

$\frac{\mu}{2} = \frac{1-\lambda_2}{2} \le h(G) \le \sqrt{2(1-\lambda_2)} = \sqrt{2\mu}$

Tight example for upper bound for Cheeger.

Eigenvalues of cycle?

Eigenvalues: $\cos\frac{2\pi k}{n}$.

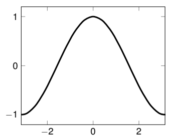
$x_i = \cos\frac{2\pi ki}{n}$

$(Mx)_i = \cos\left(\frac{2\pi k(i+1)}{n}\right) + \cos\left(\frac{2\pi k(i-1)}{n}\right) = 2\cos\left(\frac{2\pi k}{n}\right)\cos\left(\frac{2\pi ki}{n}\right)$

Cuz: $\cos(x+y) + cos(x-y) = 2\cos x \cos y$.
cuz: $e^{i(x+y)} = e^{ix}e^{iy}$. $e^{ix} = \cos x + i \sin x$. ....

Eigenvalue: $\cos\frac{2\pi k}{n}$.

Eigenvalues: vibration modes of system. Fourier basis.



$\cos x \approx 1 + \sin(0)x - \cos(0)\frac{x^2}{2}$.

$\cos\frac{2\pi}{n} \approx 1 + \sin(0)x - \cos(0)\frac{x^2}{2}$.

Second biggest eigenvalue: $\approx 1 - \frac{2\pi^2}{n^2}$.

$h(G) = \frac{2}{n} \le \frac{\pi}{n} \approx \sqrt{2\mu}$.

Random Walk.

p - probability distribution.

Probability distrubtion after choose a random neighbor.
Mp.

Converge to uniform distribution.

Power method: $M^t x$ goes to highest eigenvector.

$M^t x = a_1 \lambda_1^t v_1 + a_2 \lambda_2 v_2 + \cdots$

$\lambda_1 - \lambda_2$ - rate of convergence.

$\Omega(n^2)$ steps to get close to uniform.

Start at node 0, probability distribution, $[1, 0, 0, \cdots, 0]$.
Takes $\Omega(n^2)$ to get n steps away.

Recall drunken sailor.

Other views

Quotient Rayleigh: $\max_{x \perp 1} \frac{x^T M x}{x^t x}$.

Alternatively: $\mu = \min_{x \perp 1} \frac{x^T (L/d) x}{x^t x}$.

$L = dI - A$ or $L/d = I - A/d = I - M$.

Also:
$$x^T L x = \sum_i d x_i^2 - \sum_{e=(i,j)} 2 x_i x_j$$
$$= \sum_{e=(i,j)} (x_i^2 + x_j^2 - 2 x_i x_j)$$
$$= \sum_{e=(i,j)} (x_i - x_j)^2.$$

Warmup exercise: no larger eigenvalue on cycle.

Consider x for path.

Sort by x-value: $x_1, \ldots, x_n$.

Let $x_1^2 + x_n^2 = 1$. Shift so that $x_{n/2} = 0$.

Now, since any cut ≥ 1
$x^T L x \geq 1 \times \sum_i (x_i - x_{i+1})^2$.

Cuz: $\geq$ one edge "crossing" over the interval $[x_i, x_{i+1}]$.

Take $a = (x_i - x_{i+1})$ and $b = 1/\sqrt{n}$.

Cauchy-Schwartz $|a||b| \geq a \cdot b$.

Yields: $\sum_i (x_i - x_{i+1})^2 \geq (x_1 - x_n)^2 / n$

Furthermore: $\sum_i x_i^2 \leq n(x_1^2 + x_n^2) \leq n(x_1^2 + x_n^2 - 2x_1 x_n) = n(x_1 - x_n)^2$.

$\rightarrow \mu = \frac{x^T L x}{x^T x} \geq 1/n^2$.

Argument uses cut size to lower bound eigenvalue.
Or use eigenvalue to upper bound on cut size.
Is $\mu \geq 2/n^2$?

Interpretation of quadratic form

Consider $\mu = \min_{x \perp 1} \frac{x^T L x}{d x^T x}$.

$\sum_i x_i = 0$

Claim: $n \sum_i x_i^2 = \frac{1}{n} \sum_{i \leq j} (x_i - x_j)^2$

$\sum_i (2(n-1) x_i^2 - 2x_i \sum_{j \neq i} x_j)$.

$-2x_i \sum_{j \neq i} x_j) = -2x_i((\sum_j x_j) - x_i) = 2x_i^2 + x_i \sum_i x_j = 2x_i^2$

$\sum_i n x_i^2 - (\sum_i x_i)(\sum_j x_j)$.

Claim: $x^T L x = \sum_{e=(i,j)} M_{ij} (x_i - x_j)^2$

$\mu = \min_{x \perp 1} \frac{\sum_e M_{ij} (x_i - x_j)^2}{\frac{d}{n} \sum_{i,j} (x_i - x_j)^2}$

Ratio of average edge "length" to pair length.

Conductance $\phi(G) = n \frac{E(S, \overline{S})}{|S||V - S|}$.

Idea: pick random threshold.
Average number of edges cut compared to pairs is μ?

Random threshold.

$\mu = \min_{x \perp 1} \frac{\sum_e M_{ij} (x_i - x_j)^2}{\frac{d}{n} \sum_{i,j} (x_i - x_j)^2}$

Conductance $\phi(G) = n \frac{E(S, \overline{S})}{|S||V - S|}$.

Idea: pick random threshold.

Average number of edges cut compared to pairs is μ.
True if "length" corresponded to $|x_i - x_j|$.
Off when length is $(x_i - x_j)^2$.

Off by average length of edges.
For path:
edges have $|x_i - x_j|$ being $1/n$ of that for pairs.
For expanders:
edges go a constant fraction of what random pairs do.

Cheeger Hard Part.

Now let's get to the hard part of Cheeger $h(G) \leq \sqrt{2(1 - \lambda_2)}$.

Idea: We have $1 - \lambda_2$ as a continuous relaxation of $\phi(G)$

Take the 2^{nd} eigenvector $x = \text{argmin}_{x \in \mathbb{R}^V - \text{Span}\{\mathbf{1}\}} \frac{\sum_{i,j} M_{ij} (x_i - x_j)^2}{\frac{1}{n} \sum_{i,j} (x_i - x_j)^2}$

Consider x as an embedding of the vertices to the real line.

Round x to get a $x \in \{0, 1\}^{|V|}$

Rounding: Take a threshold t,
$$\begin{cases} x_i \geq t & \rightarrow x_i = 1 \\ x_i < t & \rightarrow x_i = 0 \end{cases}$$

What will be a good t?

We don't know. Try all possible thresholds ($n - 1$ possibilities), and hope there is a t leading to a good cut!

Sweep Cut Algorithm

Input: $G = (V, E)$, $x \in \mathbb{R}^V, x \perp \mathbf{1}$

Sort the vertices in non-decreasing order with respect to x
WLOG $V = \{1, \ldots, n\}$ $\quad x_1 \leq x_2 \leq \ldots \leq x_n$

Let $S_i = \{1, \ldots, i\}$ $\quad i = 1, \ldots, n-1$

Return $S = \text{argmin}_{S_i} h(S_i)$

Main Lemma: $G = (V, E)$, d-regular
$x \in \mathbb{R}^V, x \perp \mathbf{1}, \mu = \frac{\sum_{i,j} M_{ij} (x_i - x_j)^2}{\frac{d}{n} \sum_{i,j} (x_i - x_j)^2}$

If S is the ouput of the sweep cut algorithm, then $h(S) \leq \sqrt{2\mu}$

Note: Applying the Main Lemma with the 2^{nd} eigenvector v_2, we have $\mu = 1 - \lambda_2$, and $h(G) \leq h(S) \leq \sqrt{2(1 - \lambda_2)}$. Done!

Proof of Main Lemma

WLOG $V = \{1, \ldots, n\}$ $\quad x_1 \le x_2 \le \ldots \le x_n$

Want to show

$$\exists i \text{ s.t. } h(S_i) = \frac{\frac{1}{d}|E(S_i, V - S_i)|}{\min(|S_i|, |V - S_i|)} \le \sqrt{2\mu}$$

Probabilistic Argument: Construct a distribution D over $\{S_1, \ldots, S_{n-1}\}$ such that

$$\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S, V-S)|]}{\mathbb{E}_{S\sim D}[\min(|S|, |V-S|)]} \le \sqrt{2\mu}$$

$$\rightarrow \mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)| - \sqrt{2\mu}\min(|S|,|V-S|)] \le 0$$

$$\exists S \quad \frac{1}{d}|E(S,V-S)| - \sqrt{2\mu}\min(|S|,|V-S|) \le 0$$

The distribution D

WLOG, shift and scale so that $x_{\lfloor \frac{n}{2} \rfloor} = 0$, and $x_1^2 + x_n^2 = 1$

Take t from the range $[x_1, x_n]$ with density function $f(t) = 2|t|$.

Check: $\int_{x_1}^{x_n} f(t)\mathrm{d}t = \int_{x_1}^{0} -2t\mathrm{d}t + \int_{0}^{x_n} 2t\mathrm{d}t = x_1^2 + x_n^2 = 1$

$S = \{i : x_i \le t\}$

Let D be distribution over $S_1, \ldots, S_{n-1}$ from the above process.

Denominator.

Goal: $\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \le \sqrt{2\mu}$

Denominator:

Let $T_i =$ indicator for "i is in the smaller set of $S, V - S$"

Can check

$$\mathbb{E}_{S\sim D}[T_i] = Pr[T_i = 1] = x_i^2$$

Idea: i in smaller set if $\tau \in [0, x_i]$ or $[x_i, 0]$.

$$\begin{aligned}\mathbb{E}_{S\sim D}[min(|S|, |V-S|)] &= \mathbb{E}_{S\sim D}[\sum_i T_i] \\ &= \sum_i \mathbb{E}_{S\sim D}[T_i] \\ &= \sum_i x_i^2\end{aligned}$$

Numerator

Goal: $\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \le \sqrt{2\mu}$

Numerator:

Let $T_{i,j} =$ indicator for i, j is cut by $(S, V - S)$

$$\begin{cases} x_i, x_j \text{ same sign:} & Pr[T_{i,j} = 1] = |x_i^2 - x_j^2| \\ x_i, x_j \text{ different sign:} & Pr[T_{i,j} = 1] = x_i^2 + x_j^2 \end{cases}$$

A common upper bound: $\mathbb{E}[T_{i,j}] = Pr[T_{i,j} = 1] \le |x_i - x_j|(|x_i| + |x_j|)$

$$\begin{aligned}\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|] &= \frac{1}{2}\sum_{i,j} M_{ij}\mathbb{E}[T_{i,j}] \\ &\le \frac{1}{2}\sum_{i,j} M_{ij}|x_i - x_j|(|x_i| + |x_j|)\end{aligned}$$

Cauchy-Schwarz Inequality

$|a \cdot b| \le \|a\|\|b\|$, as $a \cdot b = \|a\|\|b\|cos(a, b)$

Applying with $a, b \in \mathbb{R}^{n^2}$ with $a_{ij} = \sqrt{M_{ij}}|x_i - x_j|, b_{ij} = \sqrt{M_{ij}}|x_i| + |x_j|$

Numerator:

$$\begin{aligned}\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|] &= \frac{1}{2}\sum_{i,j} M_{ij}\mathbb{E}[T_{i,j}] \\ &\le \frac{1}{2}\sum_{i,j} M_{ij}|x_i - x_j|(|x_i| + |x_j|) \\ &= \frac{1}{2} a \cdot b \\ &\le \frac{1}{2}\|a\|\|b\|\end{aligned}$$

Simplify numerator.

Recall $\mu = \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i - x_j)^2}, a_{ij} = \sqrt{M_{ij}}|x_i - x_j|, b_{ij} = \sqrt{M_{ij}}(|x_i| + |x_j|)$

$$\begin{aligned}\|a\|^2 &= \sum_{i,j} M_{ij}(x_i - x_j)^2 = \frac{\mu}{n}\sum_{i,j}(x_i - x_j)^2 \\ &= \frac{\mu}{n}\sum_{i,j}(x_i^2 + x_j^2) - \sum_{i,j} 2x_i x_j \\ &= \frac{\mu}{n}\sum_{i,j}(x_i^2 + x_j^2) - 2(\sum_i x_i)^2 \\ &\le \frac{\mu}{n}\sum_{i,j}(x_i^2 + x_j^2) = 2\mu\sum_i x_i^2\end{aligned}$$

$$\begin{aligned}\|b\|^2 &= \sum_{i,j} M_{ij}(|x_i| + |x_j|)^2 \\ &\le \sum_{i,j} M_{ij}(2x_i^2 + 2x_j^2) \\ &= 4\sum_i x_i^2\end{aligned}$$

Put together.

Goal: $\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \leq \sqrt{2\mu}$

Numerator:

$$\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|] = \leq \frac{1}{2}\|a\|\|b\|$$
$$\leq \frac{1}{2}\sqrt{2\mu\sum_i x_i^2}\sqrt{4\sum_i x_i^2} \quad = \sqrt{2\mu}\sum_i x_i^2$$

Recall **Denominator:**

$$\mathbb{E}_{S\sim D}[min(|S|,|V-S|)] = \sum_i x_i^2$$

We get

$$\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \leq \sqrt{2\mu}$$

Thus $\exists S_i$ such that $h(S_i) \leq \sqrt{2\mu}$, which gives $h(G) \leq \sqrt{2(1-\lambda)}$ $\square$

Kind of a proof.

$G = (V,E)$, $h = h(G)$.

Claim:
From $S \subset V$ of vertices, $|E(N_{1/h}(S))| \geq 2|E(S)|$.

Claim': there are $\Omega(h|S|)$ paths of length $\ell = 1/h$ in $N_{1/h}(S)$.

Cut size is $\geq h(G)|S| \implies$ flow of value $h(G)|S|$.
Max flow-min cut theorem.

From path argument: $\implies \mu \geq \frac{1}{\ell^2} = h(G)^2$.

Run argument over sets of size 2^i and one gets the upper bound.

Why no $\log n$ factor?
The mass splits, and **every level** has $\mu_i \geq h(G)^2$.

Cheeger proof magically does this!