Class 7

Sparsest Cuts from Metric Embeddings

Warm-Up

Group Work

Let $G = (V, E)$ be a weighted, undirected graph, on n vertices with edge weights w_{uv} on the edge $\{u, v\} \in E$. Let $d : V \times V \to \mathbb{R}$ be the associated graph metric.

Explain how to efficiently find and apply a map $f : V \to \mathbb{R}^k$, for $k = O(\log^2 n)$, so that

$$\frac{\sum_{\{u,v\}\in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\}\in\binom{V}{2}} \|f(u) - f(v)\|_1} \leq O(\log n) \frac{\sum_{\{u,v\}\in E} d(u,v)}{\sum_{\{u,v\}\in\binom{V}{2}} d(u,v)}$$

holds with high probability. Above, $\binom{V}{2}$ refers to the set of all unordered pairs $\{u, v\}$ for $u, v \in V$ and $u \neq v$.

Announcements

- HW2 due Friday!
- HW3 out now (or soon!)

Recap

- Bourgain's embedding!
  - Randomized embedding from *any* X of size n into $(\mathbb{R}^k, \ell_1)$
  - Distortion $O(\log n)$
  - $k = O(\log^2 n)$

Questions?

Minilectures, quiz, warmup?

Group Work

Let $G = (V, E)$ be a weighted, undirected graph, on n vertices with edge weights w_{uv} on the edge $\{u, v\} \in E$. Let $d : V \times V \to \mathbb{R}$ be the associated graph metric.

Explain how to efficiently find and apply a map $f : V \to \mathbb{R}^k$, for $k = O(\log^2 n)$, so that

$$\frac{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in \binom{V}{2}} \|f(u) - f(v)\|_1} \leq O(\log n) \frac{\sum_{\{u,v\} \in E} d(u,v)}{\sum_{\{u,v\} \in \binom{V}{2}} d(u,v)}$$

holds with high probability. Above, $\binom{V}{2}$ refers to the set of all unordered pairs $\{u, v\}$ for $u, v \in V$ and $u \neq v$.

Warm-Up

Let $G = (V, E)$ be a weighted, undirected graph, on n vertices with edge weights w_{uv} on the edge $\{u, v\} \in E$. Let $d : V \times V \to \mathbb{R}$ be the associated graph metric.

Explain how to efficiently find and apply a map $f : V \to \mathbb{R}^k$, for $k = O(\log^2 n)$, so that

$$\frac{\sum_{\{u,v\}\in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\}\in\binom{V}{2}} \|f(u) - f(v)\|_1} \le O(\log n) \frac{\sum_{\{u,v\}\in E} d(u,v)}{\sum_{\{u,v\}\in\binom{V}{2}} d(u,v)}$$

holds with high probability. Above, $\binom{V}{2}$ refers to the set of all unordered pairs $\{u, v\}$ for $u, v \in V$ and $u \neq v$.

- Use Bourgain's embedding!

$$\frac{k}{b \log n} d(u,v) \le ||f(u) - f(v)||_1 \le k d(u,v)$$

- Apply to top and bottom.

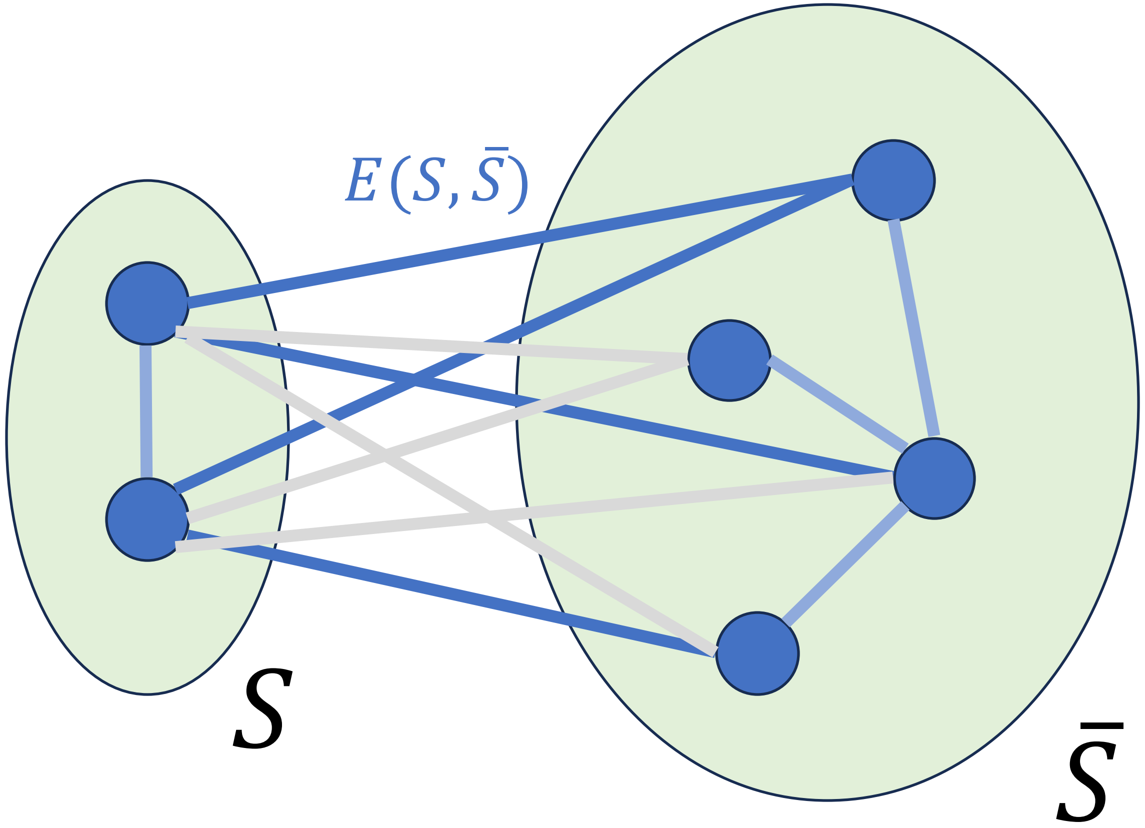
Plan for today

- Application of Bourgain’s embedding to sparsest cuts!

Sparsest Cuts

$$\phi(G,S) = \frac{|E(S,\bar{S})|}{|S||\bar{S}|}$$

$$\phi(G) = \min_S \phi(G,S)$$



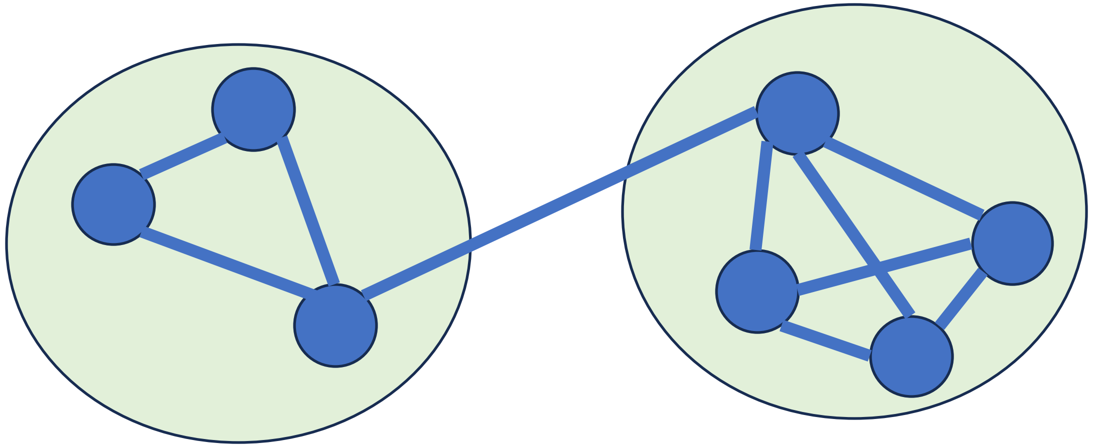
Number of possible edges between $S, \bar{S}$: $|S||\bar{S}|$

Sparsest Cut = cut that realizes $\phi(G)$

$$\phi(G,S) = \frac{|E(S,\bar{S})|}{|S||\bar{S}|}$$

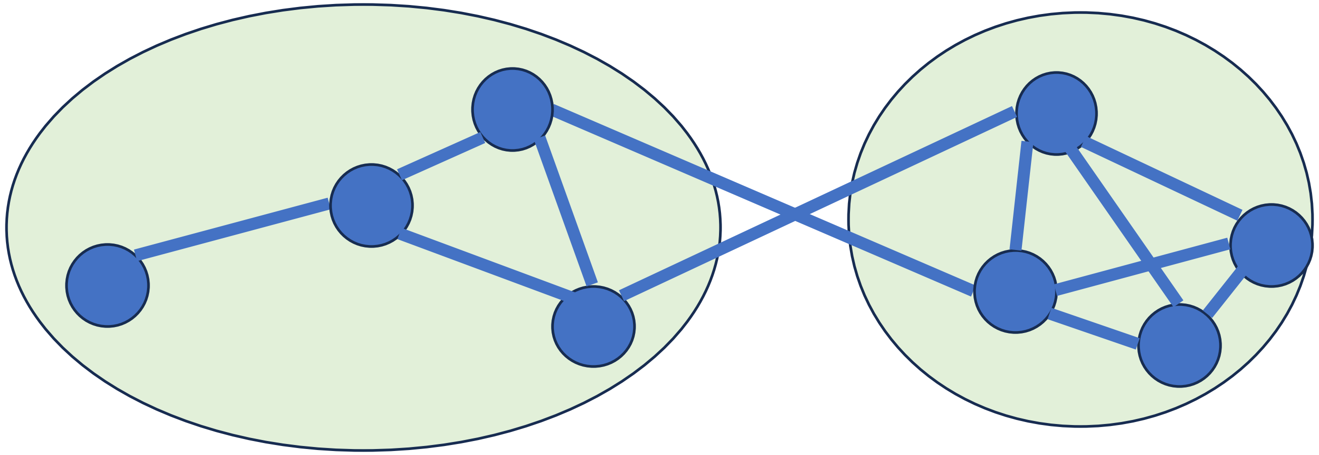
$$\phi(G) = \min_S \phi(G,S)$$

- What's the sparsest cut of this graph?



$$\phi(G) = \frac{1}{12}$$

- This one?



$$\phi(G) = \frac{2}{16} = \frac{1}{8}$$

NOT the same as a **minimum** cut!

Efficient algorithms for sparsest cut?

- We saw an algorithm for **min-cut** in Week 1!
  - Karger's algorithm!
- …But it turns out, **sparsest-cut** is NP hard!

- Today, we'll see a randomized **approximation algorithm** for sparsest cut.
  - Returns S so that $\phi(G,S) \leq O(\log n) \cdot \phi(G)$
  - (Probably)

Assuming plausible complexity-theoretic assumptions, it's NP-hard even to approximate $\phi(G)$ to within a constant factor.

So the $O(\log n)$ is pretty good!

Outline

- First group work: Show that

$$\phi(G) = \inf_{f:V\to\mathbb{R}^k} \frac{\sum_{\{u,v\}\in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\}\in\binom{V}{2}} \|f(u) - f(v)\|_1}$$

- Second group work:
  - Use this to design an approximation algorithm.

Group Work!

1. $$\phi(G) = \min_{f:V \to \{0,1\}} \frac{\sum_{\{u,v\} \in E} |f(u) - f(v)|}{\sum_{\{u,v\} \in \binom{V}{2}} |f(u) - f(v)|},$$

This one is the conceptually important one

2. $$\phi(G) = \inf_{f:V \to \mathbb{R}} \frac{\sum_{\{u,v\} \in E} |f(u) - f(v)|}{\sum_{\{u,v\} \in \binom{V}{2}} |f(u) - f(v)|},$$

3. $$\phi(G) = \min_{f:V \to \mathbb{R}^k} \frac{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in \binom{V}{2}} \|f(u) - f(v)\|_1},$$

Just try to get some intuition for these.

Solutions: Part 1, Problem 1

$$\phi(G) = \min_{f:V\to\{0,1\}} \frac{\sum_{\{u,v\}\in E} |f(u) - f(v)|}{\sum_{\{u,v\}\in\binom{V}{2}} |f(u) - f(v)|},$$

$f : V \rightarrow \{0,1\}$ ⟷ $S \subseteq V$

$$f_S(x) = 1[x \in S]$$

$$S_f = \{\, x : f(x) = 1 \,\}$$

Numerator: $\sum_{\{u,v\}\in E} |f(u) - f(v)|$ ⟷ $|E(S, \bar{S})|$

Denominator: $\sum_{\{u,v\}\in\binom{V}{2}} |f(u) - f(v)|$ ⟷ $|S||\bar{S}|$

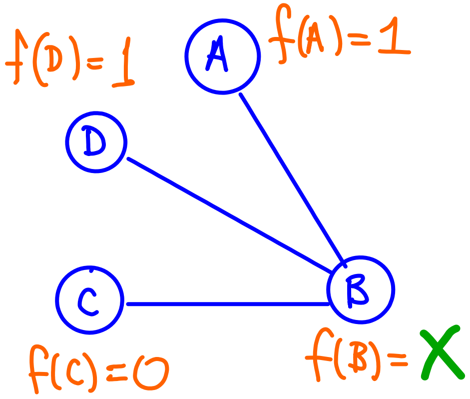
Solution: Problem 2

Note: this is just meant as intuition

$$\varphi(G) = \min_{f:V\to\mathbb{R}} \frac{\sum_{\{u,v\}\in E} |f(u)-f(v)|}{\sum_{\{u,v\}\in\binom{V}{2}} |f(u)-f(v)|}$$

Call this $R(f)$

EXAMPLE: Say $f: V \to \{0, \tfrac{1}{2}, 1\}$



$$R(f) = \frac{|1-\tfrac{1}{2}| + |\tfrac{1}{2}-0| + |\tfrac{1}{2}-1|}{|1-\tfrac{1}{2}| + |\tfrac{1}{2}-0| + |\tfrac{1}{2}-1| + |1-0| + |1-1| + |1-0|}$$

$$R(x) = \frac{|1-x| + |x-0| + |x-1|}{|1-x| + |x-0| + |x-1| + |1-0| + |1-1| + |1-0|}$$

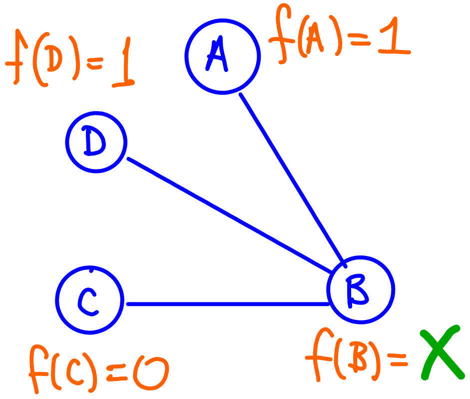
Solution: Problem 2

Note: this is just meant as intuition

$$\varphi(G) = \min_{f: V \to \mathbb{R}} \frac{\sum_{\{u,v\} \in E} |f(u) - f(v)|}{\sum_{\{u,v\} \in \binom{V}{2}} |f(u) - f(v)|}$$

Call this $R(f)$

EXAMPLE: Say $f: V \to \{0, \tfrac{1}{2}, 1\}$



$$R(x) = \frac{|1 - x| + |x - 0| + |x - 1|}{|1 - x| + |x - 0| + |x - 1| + |1 - 0| + |1 - 1| + |1 - 0|}$$

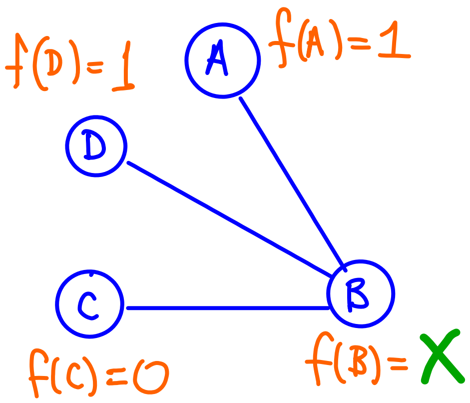
Solution: Problem 2

Note: this is just meant as intuition

$$\varphi(G) = \min_{f: V \to \mathbb{R}} \frac{\sum_{\{u,v\} \in E} |f(u) - f(v)|}{\sum_{\{u,v\} \in \binom{V}{2}} |f(u) - f(v)|}$$

Call this $R(f)$

EXAMPLE: Say $f: V \to \{0, \tfrac{1}{2}, 1\}$



$$R(x) = \frac{|1-x| + |x-0| + |x-1|}{|1-x| + |x-0| + |x-1| + |1-0| + |1-1| + |1-0|}$$

for $x \in [0,1]$...

$$R(x) = \frac{(1-x) + (x-0) + (1-x)}{(1-x) + (x-0) + (1-x) + 2} = \frac{2-x}{4-x}$$

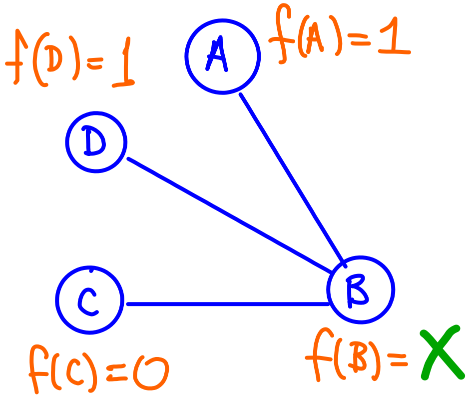
Solution: Problem 2

Note: this is just meant as intuition

$$\varphi(G) = \min_{f:V\to\mathbb{R}} \frac{\sum_{\{u,v\}\in E} |f(u)-f(v)|}{\sum_{\{u,v\}\in\binom{V}{2}} |f(u)-f(v)|}$$

Call this $R(f)$

EXAMPLE: Say $f : V \longrightarrow \{0, 1/2, 1\}$



for $x \in [0, 1]$ ... $R(x) = \dfrac{2-x}{4-x}$

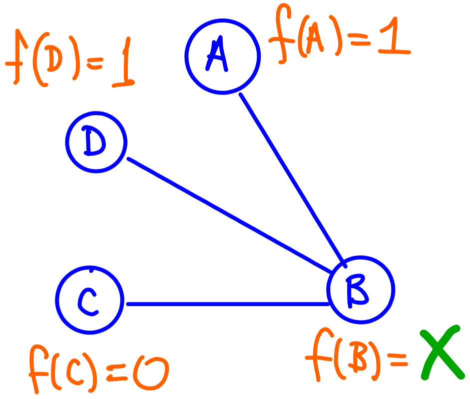
Solution: Problem 2

Note: this is just meant as intuition

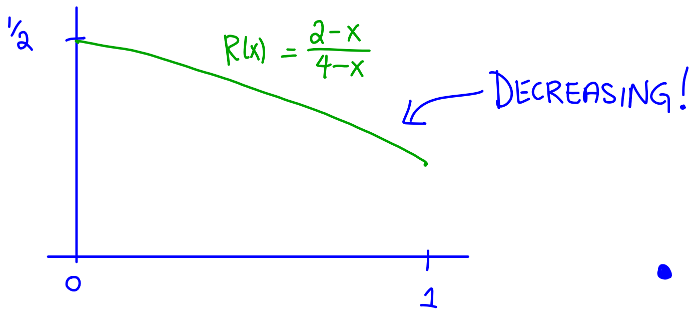
$$\varphi(G) = \min_{f: V \to \mathbb{R}} \frac{\sum_{\{u,v\} \in E} |f(u) - f(v)|}{\sum_{\{u,v\} \in \binom{V}{2}} |f(u) - f(v)|}$$

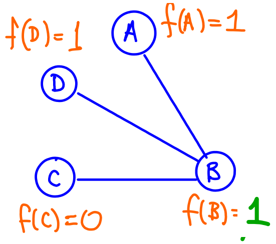
Call this $R(f)$

EXAMPLE: Say $f: V \to \{0, \frac{1}{2}, 1\}$



for $x \in [0,1]$... $R(x) = \frac{2-x}{4-x}$



- This will always be either (weakly) increasing or decreasing.
- If we replace f with , $R(f)$ doesn't increase.

- $\Rightarrow$ There is a fn $f: V \to \mathbb{R}$ that minimizes $R(f)$ that takes only two values.

Solution: Problem 2

Note: this is just meant as intuition

$$\varphi(G) = \min_{f:V\to\mathbb{R}} \frac{\sum_{\{u,v\}\in E} |f(u)-f(v)|}{\sum_{\{u,v\}\in\binom{V}{2}} |f(u)-f(v)|}$$

From before:

$$\phi(G) = \min_{f:V\to\{0,1\}} \frac{\sum_{\{u,v\}\in E} \|f(u)-f(v)\|_1}{\sum_{\{u,v\}\in\binom{V}{2}} \|f(u)-f(v)\|_1}$$

We just showed that the min over $f: V \to \mathbb{R}$ is actually attained by some $f: V \to \{0,1\}$.

So we get the same thing ($\phi(G)$) if we replace $\{0,1\}$ with $\mathbb{R}$.

Solution: Problem 3

$$\varphi(G) = \min_{f: V \to \mathbb{R}^k} \frac{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in \binom{V}{2}} \|f(u) - f(v)\|_1}$$

If $f: V \to \mathbb{R}^k$, say $f(x) = (f_1(x), \ldots, f_k(x))$,

$$\frac{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in \binom{V}{2}} \|f(u) - f(v)\|_1} = \frac{\sum_{i=1}^{k} \left(\sum_{\{u,v\} \in E} |f_i(u) - f_i(v)| \right)}{\sum_{i=1}^{k} \left(\sum_{\{u,v\} \in \binom{V}{2}} |f_i(u) - f_i(v)| \right)} \geq \min_i \underbrace{\frac{\sum_{\{u,v\} \in E} |f_i(u) - f_i(v)|}{\sum_{\{u,v\} \in \binom{V}{2}} |f_i(u) - f_i(v)|}}_{\text{this is the case where } f: V \to \mathbb{R}}$$

So adding more dimensions to f can't make this value any smaller than $f: V \to \mathbb{R}$

Conclusion

$$\phi(G) = \inf_{f:V\to\mathbb{R}^k} \frac{\sum_{\{u,v\}\in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\}\in\binom{V}{2}} \|f(u) - f(v)\|_1}$$

- Next up: using this to design an algorithm!

$$\phi(G) = \inf_{f:V\to\mathbb{R}^k} \frac{\sum_{\{u,v\}\in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\}\in\binom{V}{2}} \|f(u) - f(v)\|_1}$$

Let's come up with an algorithm!

- Hope: find f to minimize $R(f) := \frac{\sum_{\{u,v\}\in E} \|f(u)-f(v)\|_1}{\sum_{\{u,v\}\in\binom{V}{2}} \|f(u)-f(v)\|_1}$
  - Unfortunately that's not so easy…
- Instead,

Find values $d_{u,v} \in \mathbb{R}$ for all $u \neq v \in V$ to minimize

$$Q(d) := \sum_{\{u,v\}\in E} d_{u,v}$$

subject to:

- $d_{u,v} = d_{v,u} \geq 0$ for all u, v
- $d_{u,v} + d_{v,w} \geq d_{u,w}$ for all u, v, w
- $\sum_{\{u,v\}\in\binom{V}{2}} d_{u,v} = 1$

This is a **linear program.** Turns out we can solve it efficiently.

Group Work!

1. Suppose that d^* is the minimizer of the problem above.
   Explain why $Q(d^*) \leq \phi(G)$.
2. Find a randomized algorithm to approximate $\phi(G)$. More precisely, give a randomized algorithm that finds $f : V \to \mathbb{R}^k$ so that, with high probability,

$$\frac{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in \binom{V}{2}} \|f(u) - f(v)\|_1} \leq O(\log n)\phi(G).$$

3. Given f as in the previous part, explain how to efficiently find a set $S \subset V$ so that

$$\phi(G, S) \leq O(\log n)\phi(G).$$

Find values $d_{u,v} \in \mathbb{R}$ for all $u \neq v \in V$ to minimize

$$Q(d) := \sum_{\{u,v\} \in E} d_{u,v}$$

subject to:

- $d_{u,v} = d_{v,u} \geq 0$ for all u, v
- $d_{u,v} + d_{v,w} \geq d_{u,w}$ for all u, v, w
- $\sum_{\{u,v\} \in \binom{V}{2}} d_{u,v} = 1$

$$\phi(G) = \min_{f:V\to\mathbb{R}^k} \frac{\sum_{\{u,v\}\in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\}\in\binom{V}{2}} \|f(u) - f(v)\|_1}$$

Solution: Problem 1

Say d^* minimizes this LP. Show that $Q(d^*) \le \phi(G)$

- Consider d given by

$$d_{u,v} = \frac{\|f(u)-f(v)\|_1}{\sum_{\{u,v\}\in\binom{V}{2}} \|f(u)-f(v)\|_1}$$

For the f that attains the min here.

- This satisfies all the constraints.

- $$Q(d) = \frac{\sum_{\{u,v\}\in E} \|f(u)-f(v)\|_1}{\sum_{\{u,v\}\in\binom{V}{2}} \|f(u)-f(v)\|_1} = \phi(G)$$

$\rotatebox{90}{$\le$}$ (i.e. $Q(d) \ge Q(d^*)$)

$Q(d^*)$

Find values $d_{u,v} \in \mathbb{R}$ for all $u \neq v \in V$ to minimize

$$Q(d) := \sum_{\{u,v\}\in E} d_{u,v}$$

subject to:

- $d_{u,v} = d_{v,u} \ge 0$ for all u, v
- $d_{u,v} + d_{v,w} \ge d_{u,w}$ for all u, v, w
- $\sum_{\{u,v\}\in\binom{V}{2}} d_{u,v} = 1$

Solution: Problem 2

Find an algorithm! Find f s.t. $R(f) \leq O(\log n)\phi(G)$

- Solve this LP to find some d^*
- Use Bourgain's embedding to to find some $f: V \to \mathbb{R}^k$ so that

$$\frac{k}{b \log n} d^*(u,v) \leq ||f(u) - f(v)||_1 \leq k d^*(u,v)$$

- By the Warm-Up,

$$\frac{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in \binom{V}{2}} \|f(u) - f(v)\|_1} \leq O(\log n)\, Q(d^*)$$

- By previous part: $\leq O(\log n)\phi(G)$

Warm-Up:

$$\frac{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in \binom{V}{2}} \|f(u) - f(v)\|_1} \leq O(\log n) \frac{\sum_{\{u,v\} \in E} d(u,v)}{\sum_{\{u,v\} \in \binom{V}{2}} d(u,v)}$$

LP:

Find values $d_{u,v} \in \mathbb{R}$ for all $u \neq v \in V$ to minimize

$$Q(d) := \sum_{\{u,v\} \in E} d_{u,v}$$

subject to:

- $d_{u,v} = d_{v,u} \geq 0$ for all u, v
- $d_{u,v} + d_{v,w} \geq d_{u,w}$ for all u, v, w
- $\sum_{\{u,v\} \in \binom{V}{2}} d_{u,v} = 1$

Solution: Problem 3

Put it together!

- Find d^* by solving a linear program.
- Find $f: V \to \mathbb{R}^k$ via Bourgain's embedding.

We know that

$$\frac{\sum_{\{u,v\}\in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\}\in \binom{V}{2}} \|f(u) - f(v)\|_1} \leq O(\log n)\, \phi(G)$$

1. Write $f(x) = \big(f_1(x), f_2(x), \ldots, f_k(x)\big)$. Find $i^* = \mathrm{argmin_i} \frac{\sum_{\{u,v\}\in E} |f_i(u) - f_i(v)|}{\sum_{\{u,v\}\in \binom{V}{2}} |f_i(u) - f_i(v)|}$
2. While $f_{i^*}(x)$ takes on ≥ 3 distinct values $a_1 < a_2 < a_3 < \cdots$
   - Set a_2 to either a_1 or a_3, whichever makes $R(f_{i^*})$ smaller.
3. When f_{i^*} takes only two values, $a < b$, set $f_{i^*} \leftarrow \frac{f_{i^*} - a}{b - a}$
4. Let $S = \{x : f_{i^*}(x) = 1\}$ and celebrate!

Now $\phi(G, S) \leq O(\log n)\phi(G)$!

Recap

- We can find approximately-sparsest cuts efficiently!
- Step 1: $\phi(G) = \inf_{f:V \to \mathbb{R}^k} \frac{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in \binom{V}{2}} \|f(u) - f(v)\|_1}$
- Step 2: Use an LP to to find **some** metric d^* (not necessarily an ℓ_1 metric) so that this quantity is small.
- Step 3: Use Bourgain's embedding to find some f so that $||f(u) - f(v)||_1 \approx d^*(u, v)$, so that this quantity is still pretty small.
- Step 4: Reverse-engineer Step 1 to find an actual cut $S, \bar{S}$.

Next time

- More embeddings! Into ℓ_2!

- (And before next time, hand in HW2!)