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:
- HW3 due Friday!
- Please fill out feedback form
- Starting today, I’ll try to post some version of my in-class slides on the website. (Please email me or ask on Ed if I forget).
Recap

• Bourgain’s embedding!
  • Randomized embedding from any $X$ of size $n$ into $(R^k, \ell_1)$
  • 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 $u \neq v$. 
Q1 Can it be embedded?
6 Points

Consider the graph metric space \((V, d)\) induced by the following graph:

Into which space can \((V, d)\) be isometrically embedded? Select all that apply.

- \((\mathbb{R}^2, d_1)\)
- \((\mathbb{R}^2, d_2)\)
- \((\mathbb{R}^2, d_\infty)\)
Q2 An embedding

3 Points

Let \((X, d)\) be a finite metric space with \(n\) points, and write \(X = \{x_1, x_2, \ldots, x_n\}\). Consider the map \(f : X \to \mathbb{R}^n\) given by

\[
f(y) = (d(y, x_1), d(y, x_2), \ldots, d(y, x_n))
\]

Which of the following are true? Check all that apply.

- 
  \(f\) is an isometric embedding into \((\mathbb{R}^n, d_\infty)\)

- 
  \(f\) is an isometric embedding into \((\mathbb{R}^n, d_1)\)
Plan for today

• Application of Bourgain’s embedding to sparsest cuts
Sparsest Cuts

- $G = (V,E)$ is an undirected, unweighted graph:

\[
\phi(G) = \min_{S \subseteq V} \phi(G,S)
\]

\[
\phi(G,S) = \frac{|E(S, \bar{S})|}{|S| \cdot |\bar{S}|}
\]

- "Sparsity" of the cut $(S, \bar{S})$.
- Number of edges between $S$ and $\bar{S}$ in $G$.
- Number of edges between $S$ and $\bar{S}$ in the complete graph.
Goal: Find a sparsest cut

• a.k.a., find $S$ so that $\phi(G, S) = \phi(G)$
Goal: Find a sparsest cut

• a.k.a., find $S$ so that $\phi(G, S) = \phi(G)$
• Problem: this is NP-hard.

Assuming plausible complexity-theoretic assumptions, it’s NP-hard even to approximate $\phi(G)$ to within a constant factor.
Goal: Find a sparsest cut

• a.k.a., find $S$ so that $\phi(G, S) = \phi(G)$
• Problem: this is NP-hard.

• Today: randomized algorithm to (probably) find $S$ so that

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

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

• Group Work 1:
  \[ \varphi(G) = \min_{f : V \rightarrow \mathbb{R}^k} \frac{\sum_{u,v \in E} \|f(u) - f(v)\|_1}{\sum_{u,v \in E} \|f(u) - f(v)\|_1} \]

• Group Work 2:
  • ...use something about metric embeddings to approximate that thing.
Group Work!

1. \( \phi(G) = \min_{f: V \rightarrow \{0,1\}} \frac{\sum_{\{u,v\} \in E} |f(u) - f(v)|}{\sum_{\{u,v\} \in \mathcal{Y}} |f(u) - f(v)|} \)

This one is the conceptually important one

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

Just try to get some intuition for these.

3. \( \phi(G) = \min_{f: V \rightarrow \mathbb{R}^k} \frac{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in \mathcal{Y}} \|f(u) - f(v)\|_1} \)
Solution: Problem 1

\[ \varphi(G) = \min_{f: V \rightarrow \{0,1\}} \frac{\sum_{\{u,v\} \in E} |f(u) - f(v)|}{\sum_{\{u,v\} \in \mathcal{E}(G)} |f(u) - f(v)|} \]

\[ S \subseteq V \iff f: V \rightarrow \{0,1\} \]

\[ f(u) = 1_{\{u \in S\}} \text{ aka } S = \{ u \in V : f(u) = 1 \} \]

\[ \sum_{\{u,v\} \in \mathcal{E}(G)} |f(u) - f(v)| = \frac{\# \text{edges crossing the cut } S, \overline{S} \text{ in } G}{\# \text{edges crossing the cut } S, \overline{S} \text{ in } K_n} = \frac{|E(S,\overline{S})|}{|S| \cdot |\overline{S}|} = \varphi(G,S) \]
Solution: Problem 2

Note: this is just meant as intuition

**EXAMPLE:** Say \( f: V \rightarrow \{0, \frac{\sqrt{2}}{2}, 1\} \)

\[
\begin{align*}
R(f) &= \frac{|1 - \frac{\sqrt{2}}{2}| + |\frac{\sqrt{2}}{2} - 0| + |\frac{\sqrt{2}}{2} - 1|}{|1 - \frac{\sqrt{2}}{2}| + |\frac{\sqrt{2}}{2} - 0| + |\frac{\sqrt{2}}{2} - 1| + |1 - 0| + |1 - 1| + |1 - 0|} \\
\end{align*}
\]
Solution: Problem 2

Note: this is just meant as intuition

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

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

\[
R(f) = \frac{|1 - \frac{1}{2}| + |\frac{1}{2} - 0| + |\frac{1}{2} - 1|}{|1 - \frac{1}{2}| + |\frac{1}{2} - 0| + |\frac{1}{2} - 1| + |1 - 0| + |1 - 1| + |1 - 0|}
\]

\[
R(\chi) = \frac{|1 - \chi| + |\chi - 0| + |\chi - 1|}{|1 - \chi| + |\chi - 0| + |\chi - 1| + |1 - 0| + |1 - 1| + |1 - 0|}
\]
Solution: Problem 2
Note: this is just meant as intuition

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

Call this \( R(f) \)

**EXAMPLE:** Say \( f: V \to \{0, \frac{1}{2}, 1\} \)

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

\[ f(0) = 1 \quad f(A) = 1 \]
\[ f(A) = 1 \quad f(B) = 1 \]
\[ f(C) = 0 \quad f(B) = x \]
Solution: Problem 2
Note: this is just meant as intuition

Example: Say \( f: V \rightarrow \{0, v_2, 1\} \)

\[
\varphi(G) = \min_{f: V \rightarrow \mathbb{R}} \frac{\sum_{\{u,v\} \in E} |f(u) - f(v)|}{\sum_{\{u,v\} \in E} |f(u) - f(v)|}
\]

Call this \( R(f) \)

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

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

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

\[
\varphi(G) = \min_{f: V \rightarrow \mathbb{R}} \left\{ \frac{\sum_{\{u,v\} \in E} |f(u) - f(v)|}{\sum_{\{u,v\} \in E(2)} |f(u) - f(v)|} \right\}
\]

Call this $R(f)$

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

\[
R(x) = \frac{2 - x}{4 - x}
\]
Solution: Problem 2
Note: this is just meant as intuition

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

\[ f(0) = 1 \quad f(A) = 1 \]
\[ f(D) \]
\[ f(C) = 0 \quad f(B) = x \]

\[ \varphi(G) = \min_{f: V \to R} \frac{\sum_{u, v \in E} |f(u) - f(v)|}{\sum_{u, v \in \phi} |f(u) - f(v)|} \]

Call this $R(f)$

\[ R(x) = \frac{2 - x}{4 - x} \]

• This will always be either (weakly) increasing or decreasing.
Solution: Problem 2

Note: this is just meant as intuition

EXAMPLE: Say \( f : V \rightarrow \{0, \frac{1}{2}, 1\} \)

\[
\varphi(G) = \min_{f:V \rightarrow \mathbb{R}} \frac{\sum_{i \in V, j \in E} |f(i) - f(j)|}{\sum_{i \in V, j \in E} (f(i) - f(j))}
\]

Call this \( R(f) \)

\[
\begin{align*}
\text{for } x \in [0, 1] \quad R(x) &= \frac{2 - x}{4 - x} \\
\text{This will always be either (weakly) increasing or decreasing.}
\end{align*}
\]

\[
\begin{align*}
\text{If we replace } f \text{ with } \begin{cases} f(1) = 1 \\ f(2) = 1 \\ f(3) = \frac{1}{2} \\ f(4) = 0 \\ f(5) = 0 \\
\end{cases}, \quad R(f) \text{ doesn't increase.}
\end{align*}
\]

\[
\begin{align*}
\Rightarrow \text{ There is a fn } f : V \rightarrow \mathbb{R} \text{ that minimizes } R(f) \text{ that takes only two values.}
\end{align*}
\]
Solution: Problem 2
Note: this is just meant as intuition

\[ \phi(G) = \min_{f: V \to R} \frac{\sum_{u,v \in E} |f(u) - f(v)|}{\sum_{u,v \notin E} |f(u) - f(v)|} \]

• \( \Rightarrow \) There is a \( f: V \to R \) that minimizes \( R(f) \) that takes only two values.
Solution: Problem 2

Note: this is just meant as intuition

From before:

\[
\varphi(G) = \min_{f:V \to \mathbb{R}} \sum_{u,v \in E} |f(u) - f(v)| \frac{\sum_{u,v \in E} |f(u) - f(v)|}{\sum_{u,v \in (V_2)} |f(u) - f(v)|}
\]

We just showed that the min over \( f: V \to \mathbb{R} \) is actually attained by some \( f: V \to \{0,1\} \).
Solution: Problem 3

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

\[
\frac{\sum_{u,v \in E} \| f(u) - f(v) \|_1}{\sum_{u,v \in (V_2)} \| f(u) - f(v) \|_1} = \frac{\sum_{i=1}^k \left( \sum_{u,v \in E} |f_i(u) - f_i(v)| \right)}{\sum_{i=1}^k \left( \sum_{u,v \in (V_2)} |f_i(u) - f_i(v)| \right)} \geq \min_i \frac{\sum_{u,v \in E} |f_i(u) - f_i(v)|}{\sum_{u,v \in (V_2)} |f_i(u) - f_i(v)|}
\]

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

\[ \varphi(G) = \min_{f: V \to \mathbb{R}^k} \frac{\sum_{u, v \in E} \| f(u) - f(v) \|_1}{\sum_{u, v \in V_2} \| f(u) - f(v) \|_1} \]

- Next up: using this to design an algorithm!
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.
2. Suppose that $d^*$ is the minimizer of the problem above. Explain why $Q(d^*) \leq \phi(G)$.

3. Find a randomized algorithm to approximate $\phi(G)$. More precisely, give a randomized algorithm that finds $f : V \rightarrow \mathbb{R}^k$ so that, with high probability,

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

4. 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 $u,v,w$
- $\sum_{\{u,v\} \in \binom{V}{2}} d_{u,v} = 1$
Solution: Problem 2

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

\[ R(f) \]

\[ Q(d) = \sum_{\{u,v\} \in E} d_{u,v} = \frac{\sum_{\{u,v\} \in E} d_{u,v}}{\sum_{\{u,v\} \in (V,2)} d_{u,v}} \]

\[ \min \ Q(d) = \sum_{\{u,v\} \in E} d_{u,v} \]

s.t.  
- \[ d_{u,v} = d_{v,u} \geq 0 \quad \forall u,v \]
- \[ d_{u,v} + d_{v,w} \geq d_{u,w} \quad \forall u,v,w \]
- \[ \sum_{\{u,v\} \in (V,2)} d_{u,v} = 1 \]

Let \[ d_f(u,v) = \frac{\|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in (V,2)} \|f(u) - f(v)\|_1} \]

\[ Q(d_f) = R(f) \]
Solution: Problem 2

- $d_f$ is a metric, and in particular it satisfies these constraints.

- Thus, for all $f: V \rightarrow \mathbb{R}^k$, 
  $$Q(d^*) \leq Q(d_f) = R(f)$$

- $\Rightarrow Q(d^*) \leq \min_f R(f) = \phi(G)$

Let $d_f(u,v) = \frac{\|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}$

$$Q(d_f) = R(f)$$

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

s.t.
- $d_{u,v} = d_{v,u} \geq 0 \quad \forall u,v$
- $d_{u,v} + d_{v,w} \geq d_{u,w} \quad \forall u,v,w$
- $\sum_{\{u,v\} \in E} d_{u,v} = 1$
Solution: Problem 3
Find a randomized alg. to approximate $\phi(G)$

3. Find a randomized algorithm to approximate $\phi(G)$. More precisely, give a randomized algorithm that finds $f : V \rightarrow \mathbb{R}^k$ so that, with high probability,

$$\frac{\sum_{\{u,v\} \in E} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in \binom{V}{2}} \|f(u) - f(v)\|_1} \leq O(\log n) \phi(G).$$
Solution: Problem 3
Find a randomized alg. to approximate $\phi(G)$

• This was the warm-up problem!

  • Find $d^*$, and interpret it as a metric.

  • Use Bourgain's embedding on $d^*$ to obtain $f: V \rightarrow \mathbb{R}^k$
    so that $\frac{k}{\log n} d^*(u,v) \leq \|f(u) - f(v)\|_1 \leq k d^*(u,v) \ \forall u, v$.

  • Following logic from the warm-up exercise,
    $$R(f) \leq O(\log n) \ \mathbb{Q}(d^*) \leq O(\log n) \ \phi(G)$$

Technically a pseudo-metric.
Solution: Problem 4

The final algorithm

1. Find $d^*$ by solving the linear program.
2. Find $f : V \rightarrow \mathbb{R}^k$ by applying Bourgain's embedding to $d^*$
3. Write $f(x) = (f_1(x), \ldots, f_k(x))$ and find $i^* = \arg\min_i \frac{\sum_{u,v \in E} |f_i(u) - f_i(v)|}{\sum_{u,v \in V} (|f_i(u) - f_i(v)|)}$
4. While $f_i(x)$ takes on $\geq 3$ distinct values $a_1 < a_2 < a_3 < \ldots$:
   - Set $a_2$ to either $a_1$ or $a_3$, whichever makes $R(f_i)$ smaller.
5. When $f_i$ only takes 2 values, $a_1, a_2$, set $f_i \leftarrow \frac{f_i - a_1}{a_2 - a_1}$
6. Return $S = \{ u \in V : f_i(u) = 1 \}$
7. Given $f$ as in the previous part, explain how to efficiently find a set $S \subset V$ so that $\phi(G, S) \leq O(\log n)\phi(G)$. 

Recap

• We can find approximately-sparsest cuts efficiently!

• Step 1: Use an LP to find some metric $d^*$ (not necessarily an $\ell_1$ metric) so that this quantity is small.

• Step 2: Use an LP to find some metric $d^*$ (not necessarily an $\ell_1$ 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 $S, \overline{S}$. 

\[
\varphi(G) = \min_{f : V \to \mathbb{R}^k} \frac{\sum_{u,v \in E} \|f(u) - f(v)\|_1}{\sum_{u,v \in E} \|f(u) - f(v)\|_2}
\]