Metric Techniques for Approximation Algorithms **Exercises 2**
Date: August 20, 2008

1. Simple Bounds for ℓ_2.

a. Show that all 3-point metrics embed into ℓ_2 isometrically (i.e., without changing the distances).

b. Show the metric generated by the 4-node star graph $K_{1,3}$ cannot be embedded into ℓ_2 isometrically.

Hint: suppose we are given vectors x_1, x_2, x_3 representing the three leaves such that $\|x_i - x_j\| = 2$ for $1 \leq i < j \leq 3$. If y is the vector representing the center of the star, where would y be mapped such that $\|x_i - y\| = 1$ for all i?

c. Show that the 4-cycle C_4 cannot be embedded into ℓ_2 isometrically.

2. Cut Metrics and ℓ_1. Given a point set V, a cut metric δ_S corresponding to the subset $S \subseteq V$ is defined as $\delta_S(x, y) = 1$ if $|S \cap \{x, y\}| = 1$, and $\delta_S(x, y) = 0$ otherwise.

a. **(Convex Cone)** For any $D \geq 1$, if $(V, d) \overset{D}{\hookrightarrow} \ell_1$ and $c \in \mathbb{R}_{\geq 0}$ then $(V, c \times d) \overset{D}{\hookrightarrow} \ell_1$. Also, if $(V, d_1) \overset{D}{\hookrightarrow} \ell_1$ and $(V, d_2) \overset{D}{\hookrightarrow} \ell_1$ then $(V, d_1 + d_2) \overset{D}{\hookrightarrow} \ell_1$. (Hence, the set of metrics embeddable into ℓ_1 with distortion D forms a convex cone— a special case of this is the set of ℓ_1-embeddable metrics, i.e., those with $D = 1$.)

Do the set of ℓ_2-embeddable metrics form a convex cone? Prove or give a counterexample.

What about the set of ℓ_∞-embeddable metrics?

b. **(Sum of Cut Metrics)** One can show that the n-point metric (V, d) is isometrically embeddable into ℓ_1 if and only if it can be represented as a non-negative linear combination of cut metrics; that is, if and only if there are values $\alpha_S \geq 0$ for all $S \subseteq V$ such that

$$d(x, y) = \sum_{S \subseteq V} \alpha_S \cdot \delta_S(x, y).$$

(a) To show one direction, note that a single cut metric embeds into ℓ_1, and hence any non-negative linear combination of them does.

(b) For the converse, show that a 1-dimensional ℓ_1 metric (i.e., $\mathbb{R}$ equipped with the distance $d(x, y) = |x - y|$) can be written as a non-negative combination of cut metrics. (Hint: take prefixes of the line.)

Now since any n-point ℓ_1 metric d is a sum of 1-dimensional metrics, infer that it can be written as a non-negative linear combination of cut metrics.

c. We saw that because any tree metric embeds isometrically into ℓ_1, hence given any distribution $\mathcal{D}$ over tree metrics, the metric $\mathbf{E}_{T \leftarrow \mathcal{D}}[d_T(x, y)]$ embeds into ℓ_1. And the results of FRT, gave us that any n-point metric embeds into ℓ_1 with distortion $O(\log n)$.

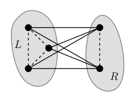
Question: Matousek showed that any n-point tree embeds into ℓ_2 with distortion $O(\sqrt{\log \log n})$. Is it possible to use this fact with the argument about to infer that general metrics embed into ℓ_2 with distortion $O(\log n \sqrt{\log \log n})$?

d. **(Cycles into ℓ_1)** Show that the (unweighted) cycles C_3 and C_4 embed into ℓ_1 isometrically. Extend the argument to show that any cycle embeds into ℓ_1 isometrically.

3. Simple Lower Bounds for ℓ_1. A metric space $\mathsf{M} = (V, d)$ satisfies the *(pure) pentagonal inequalities* if for every subset $X \subseteq V$ of five points, where $X = L \cup R$ with $|L| = 3, |R| = 2$,

$$\sum_{u,v \in L} d(u, v) + \sum_{u,v \in R} d(u, v) \leq \sum_{(u,v) \in L \times R} d(u, v).$$

I.e., the total length of the dotted edges in the figure must be at most that of the solid edges. (If $|L| = 2, |R| = 1$, you'd get the triangle inequality.)



a. Show that any n-point subset of ℓ_1 satisfies the (pure) pentagonal inequalities. Do this by showing that (a) any cut metric δ_S satisfies the inequalities, and (b) this property is maintained under taking positive linear combinations.

b. Show that the metric induced by the (unweighted) graph $K_{2,3}$ does not satisfy the inequalities. In fact, show that any embedding of $K_{2,3}$ into ℓ_1 incurs a distortion of at least 4/3.

4. Optimal Lower Bounds for ℓ_1. To show lower bounds for embedding metrics into ℓ_1, we do this: Given a metric (V, d), and non-negative values p_{ij} and q_{ij} for all $i, j \in \binom{V}{2}$, consider the bilinear form:

$$R(d) = \frac{\sum_{i,j} p_{ij}\, d(i, j)}{\sum_{i,j} q_{ij}\, d(i, j)}$$

Given a metric d, the goal is to find p, q such that for any ℓ_1-embeddable metric d', the bilinear form take very different values for $R(d)$ and $R(d')$.

a. Observe that the pentagonal-inequality lower bound can be cast in this framework: consider setting $p_{ij} = 1$ for the dotted edges, and $q_{ij} = 1$ for the solid edges.

b. We show a lower bound for embedding expander graphs into ℓ_1 using this framework. Given a constant-degree expander graph $G = (V, E)$, just set $p_{ij} = 1$ for every edge $\{i, j\} \in E$, and $q_{ij} = 1$ for all $\{i, j\} \in \binom{V}{2}$.

(a) First we show that $R(d) \le O(\frac{1}{n \log n})$.

i. Recall the fact that $\Omega(\binom{n}{2})$ pairs of vertices in the expander graph are at distance $\Omega(\log n)$ from each other.

ii. Show that $\sum_{ij} p_{ij} d_{ij} = rn/2$, where r is the degree.

iii. Also show that $\sum_{ij} q_{ij} d_{ij} = \Omega(\binom{n}{2} \log n)$.

Hence $R(d) = O(\frac{r}{n \log n})$.

(b) Now we show that $R(d') \ge \Omega(1/n)$ for every ℓ_1 metric.

i. Use problem **2(b)** to say that $d' = \sum_S \alpha_S \delta_S$ with $\alpha_S \ge 0$ (and remember that δ_S is the cut metric for set $S \subseteq V$). Hence show that

$$R(d') = \frac{\sum_{ij \in E} \sum_S \alpha_S \delta_S(i,j)}{\sum_{i,j} \sum_S \alpha_S \delta_S(i,j)}$$

ii. Now invert the order of summation, and use the fact that for $a_i, b_i \ge 0$,

$$\frac{a_1 + a_2 + \ldots + a_l}{b_1 + b_2 + \ldots + b_l} \ge \min_i \frac{a_i}{b_i}$$

to show that

$$R(d') \ge \min_{S: \alpha_S \ge 0} \frac{\sum_{ij \in E} \delta_S(i,j)}{\sum_{i,j} \delta_S(i,j)}$$

iii. Let $S^* \subseteq V$ be the cut achieving the min on the right side. Show that, without loss of generality, we can assume $|S^*| \le n/2$.

iv. Observe that for any $S \subseteq V$

$$\sum_{ij \in E} \delta_S(i,j) = |\partial S|.$$

And if $|S| \le n/2$, then

$$\sum_{ij} \delta_S(i,j) = |S| \cdot |V \setminus S| \le |S| n.$$

v. Finally, use the edge-expansion property to show that for S^*

$$|\partial S^*| / |S^*| \ge \alpha.$$

where $\alpha = \Omega(1)$ is the expansion parameter. And hence conclude that

$$R(d') \ge \frac{\alpha}{n}.$$

5. Dimension versus Distortion. The n-point uniform metric $U_n = (V, d)$ has interpoint distances $d(x,y) = 1$ for all $x \ne y \in V$.

a. Show that any embedding of U_n into $\mathbb{R}^k$ incurs a distortion of at least $\Omega(n^{1/k})$.

Hint: again, consider the vectors $x_1, x_2, \ldots, x_n$ giving the embedding. Suppose the map only expands distances, and the expansion of this map is D. What are the largest open balls around the x_i's which are disjoint? What is the smallest ball you can draw that is guaranteed to contain all the points, if the distortion is at most D?

If the volume of a radius-r ball in $\mathbb{R}^k$ is $c_k r^k$ for some constant c_k that depends only on k, what inequality does this give you?

b. Hence, if the distortion is at most $(1+\varepsilon)$, then $k \ge \frac{\log n}{\epsilon}$.

6. Trees Embed into Few Dimensions Show that any tree metric on n nodes embeds into ℓ_∞ with $O(\log n)$ dimensions.

Hint: every n-node tree has a node whose deletion breaks the tree into subtrees of size at most $2n/3$. (If you havent seen this before, it's fun to prove.)