CSC2414 - Metric Embeddings*
Lecture 13: Nonembeddability into ℓ_1

Notes taken by Hamed Hatami

Summary: In this lecture we see two nonembeddability result for ℓ_1. The first result introduces an example of a ℓ_2^2 metric which does not embed with distortion $\frac{16}{15} - \epsilon$ into ℓ_1.

The second example shows that the edit distance on the hypercube $\{0,1\}^n$ does not embed into ℓ_1 with distortion better than $\Omega(\log n)$. The proof uses the celebrated inequality of KKL.

1 Tensoring the cube

In this section we use tensoring of the cube to construct an ℓ_2^2 metric which is not ℓ_1 [HMM06]. There is an ℓ_2^2 metric space due to Khot and Vishnoi [KV05] which requires distortion $\Omega(\log\log n)$ to be embedded into ℓ_1, but the proof of that theorem is very complicated (see [KR06] for the $\Omega(\log\log n)$ bound).

For two vectors $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$, their tensor product $u \otimes v$ is a vector in $\mathbb{R}^{mn}$ defined with coordinates indexed by ordered pairs $(i,j) \in [n] \times [m]$ that assumes value $u_i v_j$ on coordinate (i,j). For example:

$$(1,2) \otimes (1,2,3) = (1,2,3,2,4,6).$$

Tensor product behaves nicely with respect to the direct product: Let $u, u' \in \mathbb{R}^n$ and $v, v' \in \mathbb{R}^n$, then

$$\langle u \otimes v, u' \otimes v' \rangle = \langle u, u' \rangle \langle v, v' \rangle. \tag{1}$$

To prove (1) note that

$$\langle u \otimes v, u' \otimes v' \rangle = \sum_{i=1}^{n} \sum_{j=1}^{m} u_i v_j u'_i v'_j = (\sum_{i=1}^{n} u_i u'_i)(\sum_{j=1}^{m} v_j v'_j) = \langle u, u' \rangle \langle v, v' \rangle.$$

Consider the hypercube $\{-1,1\}^n$, and the mapping $f : u \to u \otimes u$. Note that f maps the vertices of $\{-1,1\}^n$ to the vertices of the larger hypercube $\{-1,1\}^{n^2}$ (why?). Note that

$$\|f(u) - f(v)\|_2^2 = \langle f(u) - f(v), f(u) - f(v) \rangle = 2n^2 - 2\langle f(u), f(v) \rangle = 2n^2 - 2\langle u, v \rangle^2. \tag{2}$$

* Lecture Notes for a course given by Avner Magen, Dept. of Computer Sciecne, University of Toronto.

Since in the hypercube $\{-1,1\}^n$ the ℓ_2^2 distance is just a scaling of the ℓ_1 distance, we have that the ℓ_2^2 distance on $\{-1,1\}^n$ is in fact a metric. However this does not hold if we add the origin 0 to this set. This is because for every vector $v \in \{-1,1\}^n$, the three points v, 0, $-v$ constitute a 180 degree angle, and to have a ℓ_2^2 metric the maximum degree that we allow to have is 90. We will show that after applying the function f to the hypercube we do not face this problem anymore.

Lemma 1.1. *The set $\{f(u) : u \in \{-1,1\}^n\} \cup \{0\}$ together with the ℓ_2^2 distance constitutes a semi-metric space.*

Proof. Since $\{f(u) : u \in \{-1,1\}^n\}$ is a subset of the larger hypercube $\{-1,1\}^{n^2}$, the ℓ_2^2 distance on this set satisfies the triangle inequality. So we only need to check the triangle inequalities that involve 0. Using (2) we get

$$\|f(u)-0\|_2^2 + \|f(v)-0\|_2^2 = 2n^2 \geq \|f(u)-f(v)\|_2^2$$

and trivially

$$\|f(u)-0\|_2^2 + \|f(u)-f(v)\|_2^2 \geq \|f(v)-0\|_2^2.$$

□

The reason that in Lemma 1.1 we obtain a semi-metric instead of a metric is that f is not an injection: $f(u) = f(-u)$.

Now we show that $\{f(u) : u \in \{-1,1\}^n\} \cup \{0\}$ together with ℓ_2^2 metric does not embed well into ℓ_1. We need to use the isoperimetric inequality for the cube. Denote by Q_n the hypercube $\{-1,1\}^n$:

Theorem 1.2. *For every set $S \subseteq Q_n$,*

$$|E(S,\bar{S})| \geq |S|(n - \log_2 |S|).$$

Exercise 1.3. Use induction to prove Theorem 1.2.

Theorem 1.2 implies the following Poincaré inequality.

Proposition 1.4. *(Poincaré inequality for the cube and an additional point) Let $g : Q_n \cup \{0\} \to \ell_1$. Then the following Poincaré inequality holds.*

$$\frac{1}{2^n}\frac{16}{15}(4\alpha+1/2)\sum_{u,v\in Q_n}\|g(u)-g(v)\|_1 \leq \alpha\sum_{uv\in E}\|g(u)-g(v)\|_1 + \frac{1}{2}\sum_{u\in Q_n}\|g(u)-g(0)\|_1 \tag{3}$$

where $\alpha = \frac{\ln 2}{14-8\ln 2}$.

Proof. Let $V = Q_n \cup \{0\}$. As we have already seen many times, instead of considering $g : V \to \ell_1$ it is enough to prove the above inequality for $g : V \to \{0,1\}$. Further, we may assume without loss of generality that $g(0) = 0$. Associating S with $\{u : g(u) = 1\}$, Inequality (3) reduces to

$$\frac{1}{2^n}\frac{16}{15}(4\alpha+1/2)|S||\bar{S}| \leq \alpha|E(S,\bar{S})| + |S|/2. \tag{4}$$

From the isoperimetric inequality of Theorem 1.2 we have that $|E(S,S^c)| \geq |S|x$ for $x = n - \log_2 |S|$ and so

$$\left(\frac{\alpha x + 1/2}{1-2^{-x}}\right)\frac{1}{2^n}|S||S^c)| \leq \alpha|E(S,\bar{S})| + |S|/2.$$

It can be verified that $\frac{\alpha x+1/2}{1-2^{-x}}$ attains its minimum in $[1,\infty)$ at $x = 4$ whence $\frac{\alpha x+1/2}{1-2^{-x}} \geq \frac{4\alpha+1/2}{15/16}$, and Inequality (4) is proven. □

Theorem 1.5. *Let $V = \{u \otimes u : u \in Q_n\} \cup \{0\}$. Then for the semi-metric space X, the ℓ_2^2 metric on V, we have $c_1(X) \geq \frac{16}{15} - \epsilon$, for every $\epsilon > 0$ and sufficiently large n.*

Proof. Let $\tilde{u} = u \otimes u$. We may view X as a distance function with points in $u \in Q_n \cup \{0\}$, and $d(u,v) = \|\tilde{u} - \tilde{v}\|^2$. For every $u, v \in Q_n$, we have

$$d(u,0) = \|\tilde{u}\|^2 = \langle \tilde{u}, \tilde{u}\rangle = \langle u, u\rangle^2 = n^2,$$

and $d(u,v) = \|\tilde{u}-\tilde{v}\|^2 = \|\tilde{u}\|^2 + \|\tilde{v}\|^2 - 2\langle\tilde{u},\tilde{v}\rangle = 2n^2 - 2\langle u,v\rangle^2$. In particular, if $uv \in E$ we have $d(u,v) = 2n^2 - 2(n-2)^2 = 8(n-1)$. We next notice that

$$\sum_{u,v\in Q_n} d(u,v) = 2^{2n}\times 2n^2 - 2\sum_{u,v}\langle u,v\rangle^2 = 2^{2n}\times 2n^2 - 2\sum_{u,v}(\sum_i u_i v_i)^2 = 2^{2n}(2n^2-2n),$$

as $\sum_{u,v} u_i v_i u_j v_j$ is 2^{2n} when $i = j$, and 0 otherwise.

Let f be a nonexpanding embedding of X into ℓ_1. Using Inequality (3) we get that

$$\frac{\alpha\sum_{uv\in E}\|f(\tilde{u}) - f(\tilde{v})\|_1 + \frac{1}{2}\sum_{u\in Q_n}\|f(\tilde{v}) - f(0)\|_1}{\frac{1}{2^n}\sum_{u,v\in Q_n}\|f(\tilde{u}) - f(\tilde{v})\|_1} \geq \frac{16}{15}(4\alpha + 1/2). \qquad (5)$$

On the other hand,

$$\frac{\alpha\sum_{uv\in E} d(u,v) + \frac{1}{2}\sum_{u\in Q_n} d(u,0)}{\frac{1}{2^n}\sum_{u,v\in Q_n} d(u,v)} = \frac{8\alpha(n^2-n)+n^2}{2n^2-2n} = 4\alpha + 1/2 + o(1). \quad (6)$$

The discrepancy between (5) and (6) shows that for every $\epsilon > 0$ and for sufficiently large n, the required distortion of V into ℓ_1 is at least $16/15 - \epsilon$. □

2 Edit Distance

In this section we prove a result of Krauthgamer [KR06] that embedding the edit distance into ℓ_1 requires distortion $\Omega(\log n)$. The edit distance (a.k.a. Levenshtein distance) between two strings is the minimum number of character insertions, deletions, and substitutions needed to transform one string to the other. Let $u, v \in \{0,1\}^n$. Denote by $\mathrm{ed}(u,v)$ the edit distance between them. It is easy to see that $(\{0,1\}^n, \mathrm{ed})$ forms a metric space on the hypercube $\{0,1\}^n$.

The main tool that we use in the proof of our lower bound is an important inequality due to Kahn, Kalai, and Linial [KKL88]: For $x \in \{0,1\}^n$ and $1 \leq i \leq n$, let $x^{(i)}$ denote the vector that is the same as x except on the ith coordinate.

Theorem 2.1 (KKL Inequality). *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function with $\Pr[f(x) = 1] = p \le 1/2$, and define*

$$I_i = \Pr_x[f(x) \neq f(x^{(i)})].$$

Then

$$\max I_i \le \delta \Longrightarrow \sum_{i=1}^{n} I_i \ge \Omega(p)\log(1/\delta).$$

The highlevel view of the proof is the following: We consider f as the characteristic function of a cut. Trivially $2^n \sum I_i$ is just the number of the hypercube edges passing the cut, i.e. $E(S, \bar{S})$. When this value is small by KKL we conclude that there is one t such that I_t is large. Then because of certain symmetries on the problem we can show that there are many values of t for which I_t is large and this shows that $2^n \sum I_i$ is large which is a contradiction.

Let $V = \{0,1\}^n$ and denote by $S : V \to V$ the cyclic shift, i.e.

$$S(x_1, \ldots, x_n) = (x_n, x_1, \ldots, x_{n-1}).$$

Let

$$E = \{(x,y) : \|x - y\|_1 = 1\},$$

and

$$E_S = \{(x, S(x)) : x \in V\}.$$

We prove a Poincaré inequality:

Lemma 2.2. *Let $f : V \to \ell_1$. Then*

$$\Omega\left(\frac{\log n}{n}\right) \mathrm{avg}_{x,y\in V}\|f(x)-f(y)\|_1 \le \mathrm{avg}_{(x,y)\in E}\|f(x)-f(y)\|_1 + \mathrm{avg}_{(x,y)\in E_S}\|f(x)-f(y)\|_1.$$

Proof. We can assume that $f : V \to \{0,1\}$. Without loss of generality we can assume that $\Pr[f(x) = 1] = p \le 1/2$. Assume towards the contradiction that

$$\begin{aligned} \mathrm{avg}_{(x,y)\in E_S}\|f(x) - f(y)\|_1 &= \Pr[f(x) \neq f(S(x))] \\ &\le O(\frac{\log(n)}{n})\mathrm{avg}_{x,y\in V}\|f(x) - f(y)\|_1 \\ &\le c\frac{\log(n)}{n}p \qquad (7) \end{aligned}$$

and

$$\mathrm{avg}_{(x,y)\in E}\|f(x) - f(y)\|_1 \le c\frac{\log(n)}{n}p, \qquad (8)$$

for sufficiently small constant $c > 0$. From (7) we get that for $1 \le k \le n^{1/4}$:

$$\Pr[f(x) \neq f(S^k(x))] \leq \sum_{i=0}^{k-1} \Pr[f(S^i(x)) \neq f(S^{i+1}(x))] \leq \frac{ck \log n}{n} \leq n^{-1/2}.$$

Now notice that $(S^k(x))^{(i)} = S^k(x^{(k+i)})$. Thus for $k \leq n^{1/4}$.

$$\begin{aligned} I_j &= \Pr[f(S^k(x)) \neq f((S^k(x))^{(j)}) \\ &\leq \Pr[f(S^k(x)) \neq f(x)] + \Pr[f(x) \neq f(x^{(l+k)})] + \Pr[f(x^{(l+k)}) \neq f(S^k(x^{(l+k)}))] \\ &\leq I_{l+k} + 2n^{-1/2} \end{aligned} \tag{9}$$

Next step is to show that there exists an i such that I_i is large. Combining that with the above inequality will show that there are many values of i for which I_i is large and we get a contradiction from this. First note that

$$\sum_{i=1}^{n} I_i = n \times \text{avg}_{(x,y)\in E} \|f(x) - f(y)\|_1.$$

Thus (8) together with KKL implies that there exists some $t \in [n]$ such that

$$I_t \geq n^{-1/8}.$$

combining this with (9) we get

$$\sum_{k=1}^{n^{1/4}} I_{l+k} \geq 2n^{1/8} \geq \frac{c \log n}{n},$$

which is a contradiction. □

Now we want to use this Poincaré inequality to prove the lower bound. It is easy to see that

$$\Theta\left(\frac{1}{n}\right) \text{avg}_{x,y\in V} \text{ed}(x,y) \geq 2 \geq \text{avg}_{(x,y)\in E} \text{ed}(x,y) + \text{avg}_{(x,y)\in E_S} \text{ed}(x,y).$$

Combining this with Proposition 1.4 leads to the following theorem.

Theorem 2.3. *The edit distance on $\{0,1\}^n$ requires distortion $\Omega(\log n)$ to be embedded into ℓ_1.*

References

[HMM06] Hamed Hatami, Avner Magen, and Vangelis Markakis. Integrality gaps of semidefinite programs for vertex-cover and relations to ℓ_1 embeddability of negative type metrics. *submitted*, 2006.

[KKL88] J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. In *29-th Annual Symposium on Foundations of Computer Science*, pages 68–80, 1988.

[KR06] Robert Krauthgamer and Yuval Rabani. Improved lower bounds for embeddings into ℓ_1. In *SODA*, page to appear, 2006.

[KV05] S. Khot and N. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In *Proceedings of The 46-th Annual Symposium on Foundations of Computer Science*, 2005.