COS 521: Advanced Algorithm

Lecturer: Moses Charikar
Scribe: Tengyu Ma

1 Bourgain's Theorem

Today we are mainly going to prove the Bourgain's Theorem, which states that every metric can be embedded into ℓ_1 with logarithmic distortion. Formally,

Theorem 1 (Bourgain's Theorem)**.** *For any finite metric space (X, d) with $|X| = n$, there exists an embedding $F : X \to \mathbb{R}^m$ such that*

$$d(x, y) \le |F(x) - F(y)|_1 \le O(\log n) d(x, y)$$

To begin with, let's define a crucial notation in the proof of Bourgain's theorem

$$d(x, S) = \min_{y \in S} d(x, y)$$

Claim 2. $|d(x, S) - d(y, S)| \le d(x, y)$

Proof. WLOG, it suffices to prove that $d(x, S) - d(y, S) \le d(x, y)$. Suppose $d(y, S) = d(y, z)$ for some $z \in S$. Then $d(x, S) \le d(x, z) \le d(x, y) + d(y, z) = d(x, y) + d(y, S)$ □

The crux of the proof of the Bourgain's theorem is as follows: We want to find a distribution $\mathcal{S}$ of subsets S, such that

$$\mathbb{E}_{\mathcal{S}}[|d(x, S) - d(y, S)|] \ge \frac{c}{\log n} d(x, y)$$

Then we define the embedding as $F(x) = (d(x, S))_S$

The construction of distribution $\mathcal{S}$ is as follows:

1. Pick S_{ij}, for $j = 1, \dots, K = \log n$, and $i = 1, \dots, L$ where $L = c \log n$ for some constant c, by independently including an element of X into S_{ij} with probability $\frac{1}{2^j}$
2. Let $F(x) = (d(x, S_{11}), d(x, S_{12}), \dots, d(x, S_{\log n, L})) = (d(x, S_{ij}))_{ij}$

Thus F is an embedding into the space $\mathbb{R}^K$ where $K = L \log n = c \log^2 n$. By Claim 2 we have the following lemma:

Lemma 3.

$$|F(x) - F(y)|_1 = \sum_{ij} |d(x, S_{ij}) - d(y, S_{ij})| \le K \cdot d(x, y)$$

It suffices to bound $|F(x) - F(y)|_1$ from below by the following theorem:

Theorem 4.

$$|F(x) - F(y)|_1 \geq c_2 \log n \cdot d(x,y)$$

It can be seen that the Lemma 3 and Theorem 4 above implies that $F(x)/(c_2 \log n)$ is an embedding with distortion $O(\log n)$.

The key to prove theorem 4 is how to lower bound the value $d(x,S) - d(y,S)$ in average. The basic idea is as follows:

lower bound $d(x,S) - d(y,S)$? We draw a ball $B(x, r_1)$ with radius r_1 and center x, and a ball $B(y, r_2)$. ($B(x,r)$ is defined as $B(x,r) = \{y \in X : d(x,y) \leq r\}$). If S doesn't intersect $B(x, r_1)$ but does intersect $B(y, r_2)$, then we have that $d(x,S) \geq r_1$ and $d(y,S) \leq r_2$ and then $d(x,S) - d(y,S) \geq r_1 - r_2$. With carefully chosen r_1, r_2, we can show that the condition that "S doesn't intersect $B(x,r_1)$ but does intersect $B(y, r_2)$" happens quite often.

To see why this can be true, let assume that both $|B(x,r_1)| \approx 2^j$, $|B(y,r_2)| \approx 2^j$ for some r_1, r_2. Then we have that

$$\Pr[\ S_{ij} \text{ doesn't intersect } B(x,r_1)\ ] \approx \left(1 - \frac{1}{2^j}\right)^{2^j} \approx \frac{1}{e}$$

and

$$\Pr[\ S_{ij} \text{ intersects } B(x,r_1)\ ] \approx 1 - \left(1 - \frac{1}{2^j}\right)^{2^j} \approx 1 - \frac{1}{e}$$

This means that if we can choose $r_{1,j}$, $r_{2,j}$ such that $|B(x, r_{1,j})| \approx 2^j$, $|B(y, r_{2,j})| \approx 2^j$n hold, then we can argue that $d(x, S_{ij}) - d(x, S_{ij} \geq r_{1,j} - r_{2,j}$ □

Proof of Theorem 4. Let

$$r_j \triangleq \text{ smallest value such that } |B(x, r_j)| \geq 2^j$$

$$r'_j \triangleq \text{ smallest value such that } |B(y, r'_j)| \geq 2^j$$

and

$$\rho_j = \max\{r_j, r'_j\}$$

Let $t \triangleq$ smallest value such that $\rho_t \geq \frac{1}{2} d(x,y)$. Intuitively, when $j \geq t$, we cannot even guarantee that $B(x, r_j)$ doesn't intersect $B(y, r'_j)$, so we don't bother considering those large j's larger than t. Also, for the same technical reason, if $\rho_{t-1} + \rho_t \geq d(x,y)$, we have to redefine $\rho_t \triangleq d(x,y) - \rho_{t-1}$ so that we can guarantee that the ball $B(x, \rho_t)$ and $B(x, \rho_{t-1})$ don't intersect.

If $\rho_j = r_j$, then we define $B_j = B(x, \rho_j)$ and $G_j = B(y, \rho_{j-1})$. Note that $|B_j| = B(x, r_j) = 2^j$ and $|G_j| \geq |B(y, r'_{j-1})| = 2^{j-1}$(If $j = t$, then $|B_j| = B(x, r_j) \leq 2^j$, and the same result follows). Then if we use ρ_j, ρ_{j-1} as r_1, r_2 in the previous discussion of lower bound, we know that

$$\Pr\left[B_j = B(x, \rho_j) \text{ doesn't intersects } S_{ij}\ \right] = \left(1 - \frac{1}{2^j}\right)^{2^j} \geq \frac{1}{4}$$

and

$$\Pr\left[G_j = B(y, \rho_{j-1}) \text{ intersects } S_{ij}\ \right] \geq 1 - \left(1 - \frac{1}{2^j}\right)^{2^{j-1}} \geq \frac{1}{4}$$

Thus with probability at least $\frac{1}{16}$, we have that $d(x, S_{ij}) \geq \rho_j$ and $d(y, S_{ij}) \leq \rho_{j-1}$ and then $|d(x, S_{ij}) - d(y, S_{ij})| \geq \rho_j - \rho_{j-1}$.

On the other hand, if $\rho_j = r'_j$, we can define $B_j = B(y, \rho_j)$ and $G_j = B(x, \rho_{j-1})$, and the same argument as in previous case follows exactly. Thus we have that $|d(x, S_{ij}) - d(y, S_{ij})| \geq \rho_j - \rho_{j-1}$ with probability at least $\frac{1}{16}$. It follows that

$$\mathbb{E}[\sum_i |d(x, S_{ij}) - d(y, S_{ij})|] \geq \frac{L}{16}(\rho_j - \rho_{j-1})$$

Since for each i, S_{ij} are drawn independently, thus by Chernoff inequality, by taking $L = O(\log n)$, we have that with probability $1/2$, for some constant c_3, for any j, x, y

$$\sum_i |d(x, S_{ij}) - d(y, S_{ij})| \geq \frac{c_3 L}{16}(\rho_j - \rho_{j-1})$$

Taking sum over all j we have that

$$\sum_i \sum_j |d(x, S_{ij}) - d(y, S_{ij})| \geq \frac{c_3 L}{16}\rho_t \geq \frac{c_3 L}{16}\frac{d(x, y)}{2}$$

□

2 Other metrics embedding

Recall that in 5th Problem in the first homework, we are asked to design approximation algorithm for the k-server problem when all the servers on a line. Consider the following variants: What if all the servers are located on a circle. A idea by Karp to conquer this problem is to cut the circle at some random point, and then treat circle as a line. Thus we introduce a new metric $d_{OL}(\cdot, \cdot)$ instead of the original metric $d_C(\cdot, \cdot)$. And it can be proved that $d_{OL}(x, y) \geq d_C(x, y)$, and on the other hand, if D is the total length of the circle

$$\mathbb{E}[d_{OL}(x, y)] \leq \frac{d_C(x, y)}{D} D + \frac{1 - d_C(x, y)}{D} d_C(x, y) \leq 2 d_C(x, y)$$

Though this 2 distortion embedding in expectation sense, it suffices for the online server problem because if the algorithm on a line is α-competitive, the same algorithm is 2α-competitive on the circle.

We will talk more about metrics imbedding into tree metrics next time.