Note Title 4/18/2007

COMP 670P

Embedding arbitrary metrics into ℓ_p^d

dimension

Metric space : (V, ρ) $\rho : V \times V \to \mathbb{R}$

1. $\rho(x,y) \geq 0$
2. $\rho(x,y) = 0$ iff $x = y$
3. $\rho(x,y) = \rho(y,x)$
4. $\rho(x,z) \leq \rho(x,y) + \rho(y,z)$

$\ell_p^d : (V, \rho)$ $V = \mathbb{R}^d$ $\rho(x,y) = \|x-y\|_p$

$$\|x\|_\infty = \max_{1 \leq i \leq d} |x_i|$$

Embeddings of (V, ρ) into ℓ_p^d $(|V| = n)$

→ isometric embedding into ℓ_∞^n (Fréchet)

→ distortion $O(\log n)$, $d = O((\log n)^2)$, ℓ_∞ (Matousek)

→ distortion $O(\log n)$, $d = 2^n$, ℓ_2 (Bourgain)

JL

$d = O(\log n)$

ℓ_2

generalize to ℓ_p

$d = O((\log n)^2)$, ℓ_p

Embed (V, ρ) into ℓ_p^d

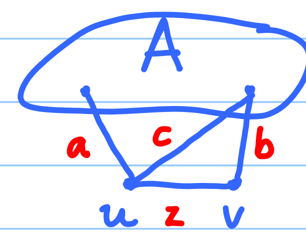
- $\phi(u) = \langle f_1(u) \cdots f_d(u) \rangle$

  D-embedding: $\frac{1}{D}\rho(u,v) \le \|\phi(u) - \phi(v)\| \le \rho(u,v)$

- choose $A_1, \ldots, A_d \subseteq V$

- set i^{th} coordinate of $u \equiv f_i(u) = \rho(u, A_i)$

• $|\rho(u, A_i) - \rho(v, A_i)| \le \rho(u, v)$

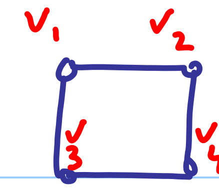


if $b + z < a$
$b + z < c$ $\Rightarrow\Leftarrow$

• $\|\phi(u) - \phi(v)\|_p \le d^{1/p} \rho(u,v)$

1 if $p = \infty$

Embedding into ℓ_∞^n



$$V = \{v_1, \dots, v_n\}$$

$$A_i = \{v_i\}$$

$$f_i(u) = \rho(u, v_i)$$

$$\|\phi(u) - \phi(v)\|_\infty \leq \rho(u,v)$$

$$f_i(v_j) - f_i(v_i) = \rho(v_i, v_j)$$

$$\Rightarrow \|\phi(u) - \phi(v)\|_\infty \geq \rho(u,v)$$

$$\begin{array}{c c} v_1 - & (0 \;\; 1 \;\; 1 \;\; 2) \\ v_2 - & (1 \;\; 0 \;\; 2 \;\; 1) \\ v_3 - & (1 \;\; 2 \;\; 0 \;\; 1) \\ v_4 - & (2 \;\; 1 \;\; 1 \;\; 0) \end{array}$$

$\ell_\infty \approx$ general metric space

* delete 1 coordinate
* $\Omega(n)$ dimension needed for 3-embedding

Given (V, ρ)

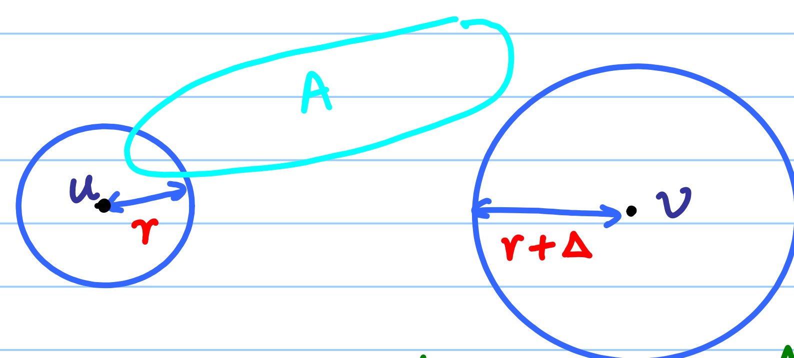
Thm: Let n be a power of 2

Let $D = 2 \log n - 1$

$\exists$ D-embedding of V into ℓ_∞^d, for $d = O((\log n)^2)$

Proof. Choose subsets of V etc. $\|\phi(u) - \phi(v)\|_\infty \leq \rho(u,v)$

want: $\forall\, u, v,\ \exists i\ |f_i(u) - f_i(v)| \geq \frac{1}{D} \rho(u,v)$



If $\exists$ such an A then

$|\rho(u,A) - \rho(v,A)| \geq \Delta$

▷ Pick enough random subsets A

$p_1 = 1/2,\ p_2 = 1/4,\ \ldots,\ p_{\log n} = 1/n$ $\qquad m = 48 \ln n$

A_{ij} — i: $1, \ldots, m$; j: $1, \ldots, \log n$

$\Pr[u \in V \text{ belongs to } A_{ij}] = p_j$ (independent)

A_{*i} : m sets of size $\sim n/2^i$

LEMMA $\forall u, v \in V\ \exists j$ s.t with prob $\geq 1/24$

$$|\rho(u, A_{ij}) - \rho(v, A_{ij})| \geq \frac{1}{D} \rho(u,v)$$

REST OF PROOF Pick j as above

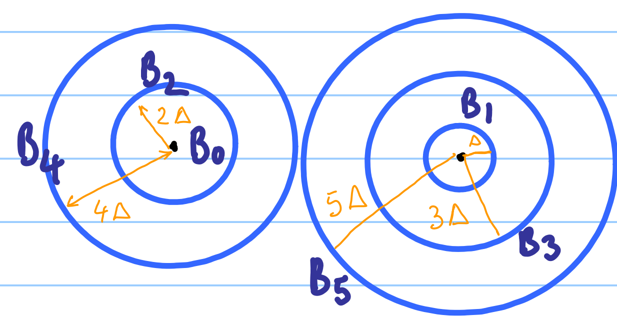
$$\left(1 - \frac{1}{24}\right)^m \leq e^{m/24} \leq n^{-2}$$

LEMMA $\forall u,v \in V \; \exists j$ s.t with prob $\geq 1/24$

$$|\rho(u, A_{ij}) - \rho(v, A_{ij})| \geq \frac{1}{D}\rho(u,v)$$

Proof: $[D = 2\log n - 1]$ $\quad \Delta = \frac{1}{D}\rho(u,v)$

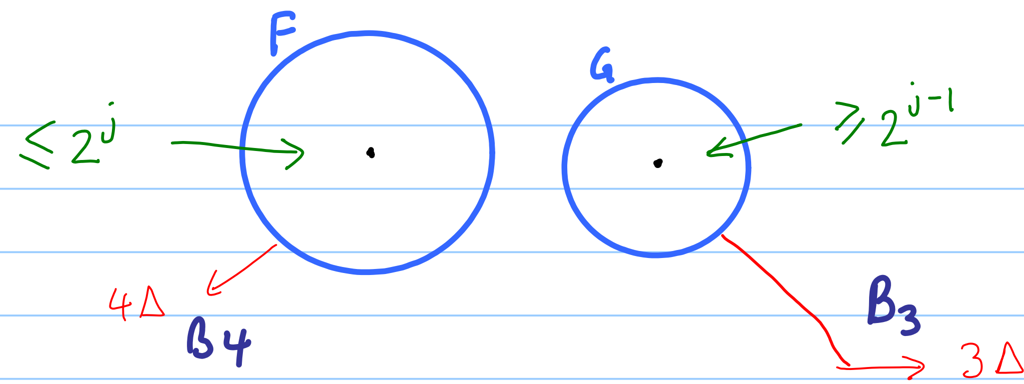
n_t = # points in B_t



Want j, t s.t

$n_t \geq 2^{j-1}$

$n_{t+1} \leq 2^j$

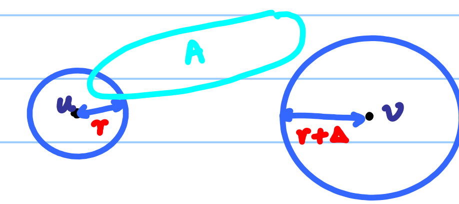


$n_t = \#$ points in B_t

j, t s.t

$n_t \geq 2^{j-1}$, $n_{t+1} \leq 2^j$

Want: random set A to hit G, avoid F.



Take A_{*j}

$\Pr[\text{Hit } G] \geq 1 - (1 - 2^{-j})^{2^{j-1}} \geq 1 - e^{-1/2} \geq 1/6$

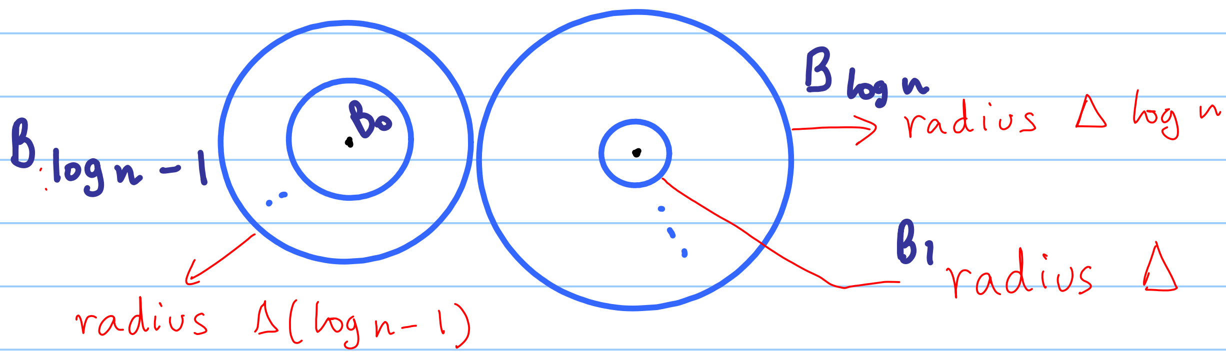
$\Pr[\text{Avoid } F] \geq (1 - 2^{-j})^{2^j} \geq 1/4$

$\Pr[\text{Avoid } F \,\&\, \text{Hit } G] \geq 1/24$

Proof: $[D = 2\log n - 1]$ $\quad \Delta = \frac{1}{D} \rho(u,v)$

$n_0 = 1 \quad n_t = \#$ points in $B_t \quad \#$ balls $= \log n + 1$

Want: $j, t,$ s.t. $n_t \geq 2^{j-1}$, $n_{t+1} \leq 2^j$



$[1,2][2,4] \cdots [n/2, n] \rightarrow \log n$ intervals

if $\exists t \quad n_t \geq n_{t+1}$ then done else

$n_0 < n_1 < n_2 \cdots < n_{\log n}$ then done by pigeonhole principle. $\blacksquare$

Overview of proof for ℓ_∞

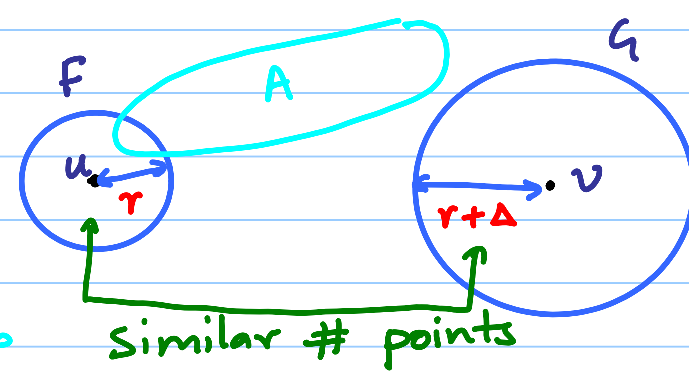
→ Choose d random subsets of V $A_1 \ldots A_d$

→ $\phi(u) = \langle \rho(u, A_1), \ldots, \rho(u, A_d) \rangle$ (nonexpanding)

→ $\forall u, v$, want index $i \in \{1, \ldots d\}$ to "take care of" putting u, v "far apart"

$$|\rho(u, A_i) - \rho(v, A_i)| \geq \frac{\rho(u,v)}{D}$$

→ Find r s.t.



random A with right density avoids G and hits F with const prob

Embedding into ℓ_2 (Bourgain's theorem)

Defn (V, ρ) (V, μ)

$\rho \geq \mu \stackrel{\text{def}}{\equiv} \forall u, v \in V \quad \rho(u,v) \geq \mu(u,v)$

Defn Line pseudometric $\varphi : V \to \mathbb{R}$

$(V, \mu) \qquad \mu(u,v) = |\varphi(u) - \varphi(v)|$

pseudometric: $\mu(u,v) = 0 \not\Rightarrow u = v$

$A \subseteq V \qquad \mu_A$ assoc. with $v \mapsto \rho(v, A)$

Claim $\forall A \subseteq V \quad \mu_A \leq \rho$

Lemma: Given (V,ρ), line p.m. $\mu_1 \dots \mu_N$ s.t

$\forall i \quad \mu_i \le \rho$ and $\sum_{i=1}^{N} \alpha_i \mu_i \ge \rho/D$ $(\alpha_i \ge 0;\ \Sigma\alpha_i = 1)$.

Then (V,ρ) can be D-embedded into ℓ_2^N

Proof Let φ_i induce μ_i

$$f : u \mapsto \langle \sqrt{\alpha_1}\,\varphi_1(u), \dots, \sqrt{\alpha_N}\,\varphi_N(u) \rangle$$

1. $$\|f(u) - f(v)\|^2 = \sum \alpha_i\, \mu_i(u,v)^2 \le \rho(u,v)^2$$

2. $$\|f(u) - f(v)\| = \left(\sum \alpha_i\, \mu_i(u,v)^2\right)^{1/2} \cdot \left(\sum \alpha_i\right)^{1/2}$$
$$\ge \sum \alpha_i\, \mu_i(u,v) \ge \rho(u,v)/D \qquad \blacksquare$$

$q = \log n$

Lemma Given $u, v \in V$ $\exists\, \Delta_1 \dots \Delta_q \geq 0$ $\quad \sum \Delta_i = \frac{1}{4} \varphi(u,v)$

s.t. $\forall j$ A_j = random subset of V $\quad \Pr[v \in A_j] = 2^{-j}$ (independent)

$$\underbrace{\Pr\Big[\ \underbrace{|\varphi(u, A_j) - \varphi(v, A_j)|}_{\mu_{A_j}} \geq \Delta_j\ \Big] \geq 1/12}$$

$$\forall u, v\ \forall j \quad \sum_{A \subseteq V} \Pr_j[A]\, \mu_A(u,v) \geq \Delta_j / 12$$

$$\sum_{j=1}^{q} \sum_{A \subseteq V} \Pr_j[A]\, \mu_A(u,v) \geq \varphi(u,v)/48$$

$$\sum_{A \subseteq V} \left(\frac{\sum_{j=1}^{q} \Pr_j[A]}{q} \right) \mu_A(u,v) \geq \varphi(u,v)/48q$$

$q = \log n$

Lemma Given $u, v \in V$ $\exists\, \Delta_1 \dots \Delta_q \geq 0$ $\quad \sum \Delta_i = \frac{1}{4}\rho(u,v)$

s.t. A_j = random subset of V $\quad \Pr[v \in A_j] = 2^{-j}$ (independent)

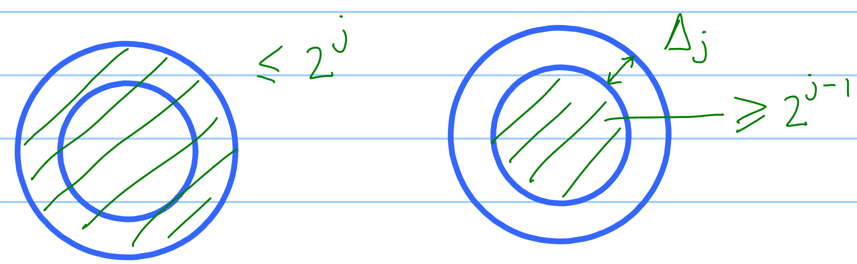
$$\Pr\left[\, |\rho(u, A_j) - \rho(v, A_j)| \geq \Delta_j \,\right] \geq 1/12$$

Proof: $r_0 = 0$

$$r_i = \min\{x \mid B(u,x) \geq 2^i \text{ and } B(v,x) \geq 2^i\}$$

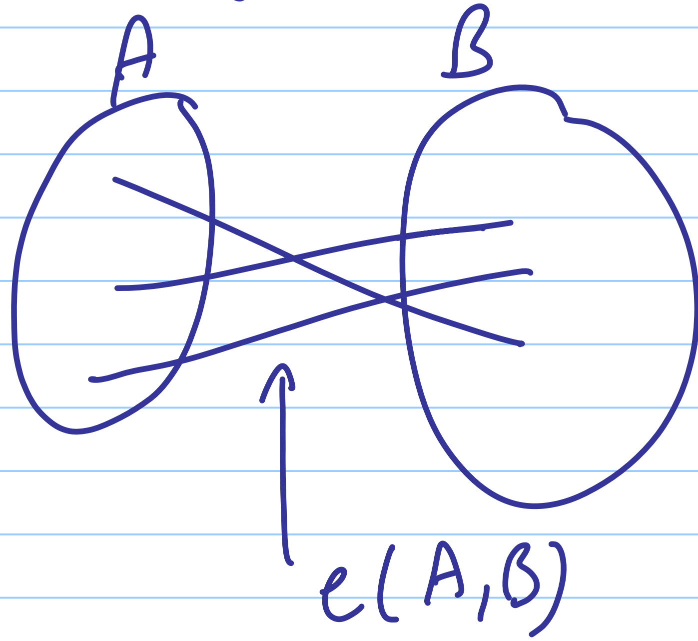
$$\underbrace{r_0 \quad r_1}_{\Delta_1} \;\underbrace{\quad}_{\Delta_2} \;\cdots\; r_k \quad \underbrace{r_{k+1} \quad \cdots \quad r_q}_{\Delta_i = 0}$$

$\rho(u,v)/4$



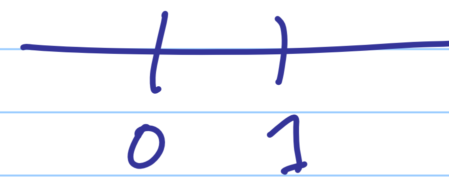
Result $O(\log n)$ approximation algorithm

min-density cut $G=(V,E)$



$$\frac{e(A,B)}{|A|\,|B|}$$

Cut pseudometric $V \to \{0, 1\}$



density of cut:

τ

A B

$$= \frac{\tau(E)}{\tau(F)}$$

F = all pairs from V

ρ : a pseudometric minimizing $\dfrac{\rho(E)}{\rho(F)}$

not a cut pseudometric

minimize $\rho(E)$
subject to $\rho(F) = 1$ } Linear prog.

Vars: ρ_{uv}
Cons: Δ ineq etc.

$(V, \rho_0) \xhookrightarrow{O(\log n)} \ell_1^{d}$ ← $\cancel{2^n}$ $O((\log n)^2)$

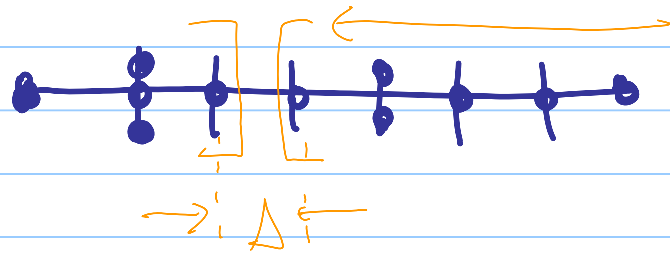
$\hookrightarrow$ new metric σ_0

d line pseudometrics ⟸

$\downarrow$

$n-1$ cut pseudometrics

Each LPM ⟺ Convex comb. of cut pseudometrics

Line PM  cut PM. Weight: Δ [Zhang]

$$\sigma_0 = \sum_i \alpha_i \gamma_i \qquad \sum_i \alpha_i = d$$

$\uparrow$ cut PMs $\qquad \alpha_i \geq 0$

$$R_1(\rho) \stackrel{\text{def}}{\equiv} \frac{\rho(E)}{\rho(F)} \qquad \forall \text{ (pseudo) metrics } \rho$$

$$R_1(\rho_0) \leq R_1(\tau_{opt}) \qquad \text{optimal cut pseudometric}$$

$$R_1(\sigma_0) \leq D \cdot R_1(\rho_0) \qquad D = O(\log n) \text{ (embedding)}$$

$$R_1(\sigma_0) = \frac{\sigma_0(E)}{\sigma_0(F)} = \frac{\sum_i \alpha_i \tau_i(E)}{\sum_i \alpha_i \tau_i(F)} \geq \min_i \frac{\tau_i(E)}{\tau_i(F)}$$

Pick i s.t. RHS minimized