Low-Distortion Embeddings II

Piotr Indyk

In the previous episode

- Definition of embedding $f:M \rightarrow M'$ with distortion c
- Isometric embedding of l_1^d into $l_\infty^{2^d}$
  - $l_\infty^{d'}$ diameter in $O(nd')$ time
  - l_1^d diameter in $O(n2^d)$ time
- Mentioned $(1+\varepsilon)$-distortion embedding of l_2^d into $l_1^{d'}$, where $d'=O(d/\varepsilon^2 \log(1/\varepsilon))$
  - $f(u) = Au$, where A is a "random" matrix
- Embedding of $M=(X,D)$ into l_∞^d
  - Isometry for $d=|X|$
  - Distortion $O(c)$ for $d=|X|^{1/c}$

Today

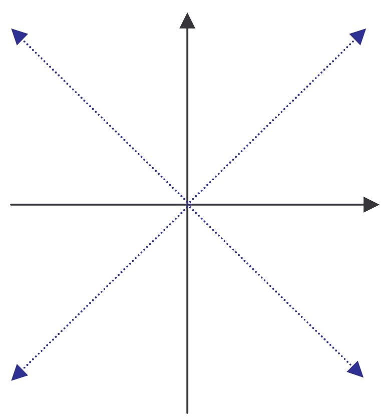
- $(1+\varepsilon)$-distortion embedding of l_2 into l_∞
  - Approximate diameter in l_2
- $(1+\varepsilon)$-distortion embedding of (X,l_2) , X in R^d, into $l_2^{d'}$, where $d'=O(\log |X|/\varepsilon^2)$
  - $(1+\varepsilon)$-approximate Near Neighbor in l_2^d
  - Query time: $O(d \log n / \varepsilon\text{^}2)$
  - Space: $n^{O(\log(1/ \varepsilon)/\varepsilon\text{^}2)}$

$(1+\varepsilon)$-embedding of l_2 into l_∞

- We know:
  - $(1+\varepsilon)$-embedding of l_2^d into $l_1^{O(d/\varepsilon^2 \log(1/\varepsilon))}$
  - Isometry of l_1^d into $l_\infty^{2^d}$
  - Therefore: $(1+\varepsilon)$-embedding of l_2^d into $l_\infty^{d'}$, where $d'=2^{O(d/\varepsilon^2 \log(1/\varepsilon))}$
- We will improve d' to $O(1/\varepsilon)^{(d-1)/2}$

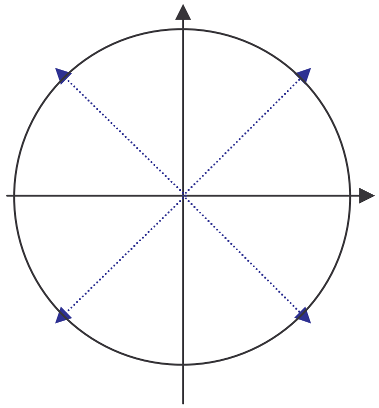
Consider $d=2$

- For embedding into l_1 we used $f(x,y)=[x+y,x-y,-x+y,-x-y]$
  - Since f linear, we have $||f(p)-f(q)||=||f(p-q)||$
  - $||(x,y)||_1 = |x|+|y| = \max[\ x+y\ ,\ x-y\ ,\ -x+y\ ,\ -x-y\ ]$

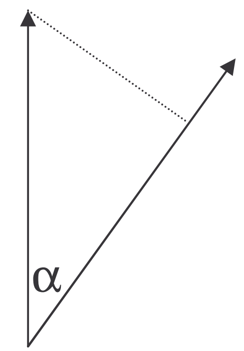


Embedding of l_2

- Again, use projections
  - Onto unit (l_2) vectors $v_1 \ldots v_k$
  - Requirement: vectors are "densely" spaced
  - I.e., for any u there is v_i such that $u*v_i \geq ||u||_2 / (1+\varepsilon)$
  - Can assume $||u||_2=1$
- How big is k ?



Lemma



- Consider two unit vectors u and v, such that the angle(u,v)=α.
  Then $u*v \geq 1-O(\alpha^2)$
- Proof: $u*v=\cos(\alpha)$

  $\approx 1+ \alpha \cos'(\alpha) + \alpha^2 \cos''(\alpha)/2$

  $\approx 1- \alpha^2/2$
- Therefore, suffices to use $2\pi/\varepsilon^{1/2}$ vectors to get distortion $1+O(\varepsilon)$

Higher Dimensions

- For $d=2$ we get $d'=O(1/\varepsilon^{1/2})$
- For any d we get $d'=O(1/\varepsilon)^{(d-1)/2}$
  - Can "cover" a unit sphere in R^d with $O(1/\alpha)^{d-1}$ vectors so that any v has angle $<\alpha$ with at least one of the vectors
  - The remainder is the same
- Yields an $O(1/\varepsilon)^{(d-1)/2}n$ – time algorithm for approximate diameter in l_2

Dimensionality Reduction

[Johnson-Lindenstrauss'85]: For any X in R^d, $|X|=n$, there is a $(1+\varepsilon)$-distortion embedding of (X,l_2), into $l_2^{d'}$, where $d'=O(\log n / \varepsilon^2)$

Proof

- Need to show that for any vector p,q in X, we have $||f(p)-f(q)|| \approx S\,||p-q||$
- Our mapping: $f(u)=Au$, A "random"
- Sufficient to show that for a *fixed* $u=p-q$, where p,q in X, we have $||Au|| \approx S||u||$ with probability at least $1-1/n^2$
- In fact, by linearity of A we can assume $||u||=1$, so we just need to show $||Au|| \approx S$

Normal Distribution

- Normal distribution:
  - Range: $(-\infty, \infty)$
  - Density: $f(x)=e^{-x^2/2} / (2\pi)^{1/2}$
  - Mean=0, Variance=1
  - If X and Y independent r.v. with normal distribution, then X+Y has normal distribution
- Basic facts:
  - $Var(cX)=c^2\ Var(X)$
  - If X,Y independent, then $Var(X+Y)=Var(X)+Var(Y)$

Back to embedding

- We map $f(u)=Au=[a^1*u,\ldots,a^{d'}*u]$, where each entry of A has normal distribution
- Consider $Z=a^i*u = a*u=\sum_i a_i u_i$
- Each term $a_i u_i$
  – Has normal distribution
  – With variance u_i^2
- Thus, Z has normal distribution with variance $\sum_i u_i^2 =1$
- This holds for each a^j

What is $||Au||_2$

- $||Au||^2 = (a^1 * u)^2 + \ldots + (a^{d'} * u)^2 = Z_1^2 + \ldots + Z_{d'}^2$
  where:
  - All Z_i's are independent
  - Each has normal distribution with variance=1
- Therefore, $E[\ ||Au||^2\ ] = d' * E[Z_1^2] = d'$
- By Chernoff-like bound

$$\Pr[\ |\ ||Au||^2 - d'| > \varepsilon d'] < e^{-B\ d'\ \varepsilon\hat{}2} < 1/n^2$$

  for some constant B

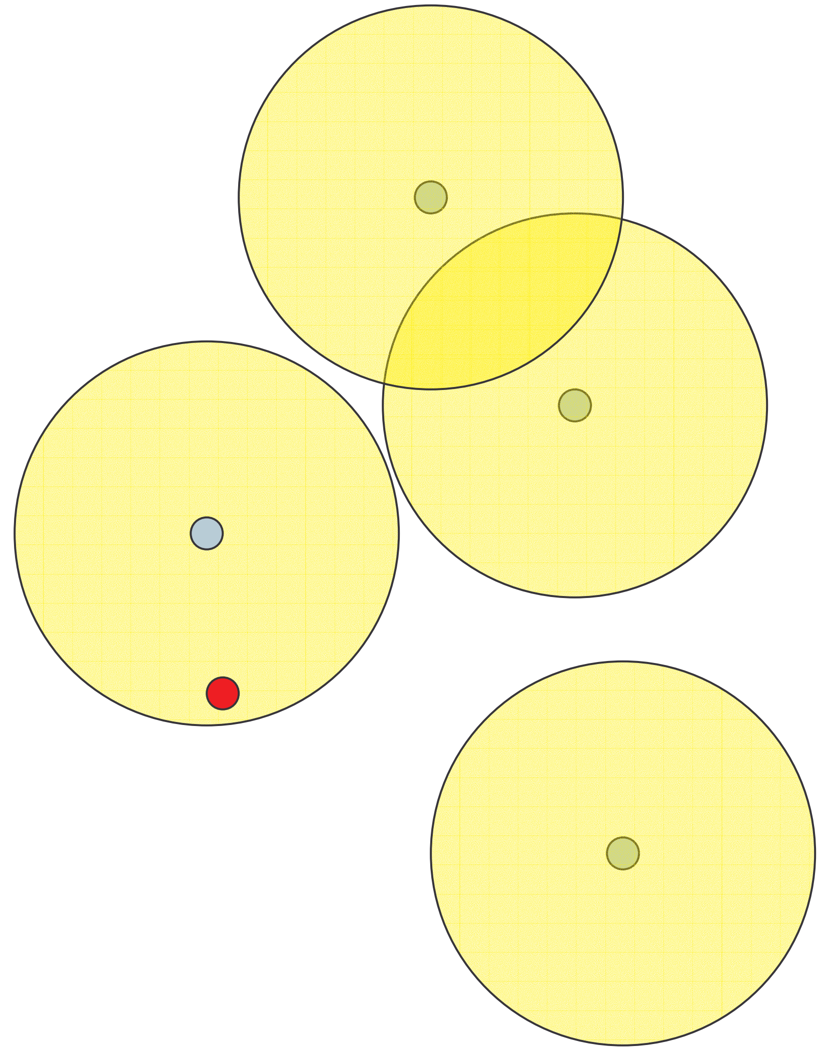
- So, $||Au||_2 \approx (d')^{1/2}$ with probability $1-1/n^2$

Application to Near Neighbor

- Suppose we have an algorithm with:
  - $O(d)$ query time
  - $O(1/\varepsilon)^d$ n space
- Then we get:
  - $O(d \log n / \varepsilon^2)$ query time
  - $n^{O(\log(1/\varepsilon)/\varepsilon^2)}$ space

$O(1/\varepsilon)^d$ n space NN

- Assume r=1



Grid

- Impose a grid with side length=$\varepsilon/d^{1/2}$
- Parameters:
  - Cell diameter: ε
  - #cells/ball: $O(1/\varepsilon)^d$
- Store all cells touching a ball