Geometric Constraints CIS 6930 Fall 2006

Lecture :Johnson-Lindenstrauss Lemma and Embedding the diamond graph in L_p

Lecturer: Dr. Meera Sitharam Scribe: Heping Gao

1 Embedding the diamond graph in L_p and dimension reduction in L_1

Definition 1 *G_0 consists of a single edge of length 1. G_i is obtained from G_{i-1} as follows. Given an edge $(u,v) \in E_{G_{i-1}}$, it is replaced by a quadrilateral u,a,v,b with edge lengths 2^{-i}. In what follows, (u,v) is called an edge of level $i-1$, and (a,b) is called the level i anti-edge corresponding to (u,v).*

Lemma 2 *Fix $1 < p \leq 2$ and $x,y,z,w \in L_p$. Then, $\|y-z\|_p^2+(p-1)\|x-w\|_p^2 \leq \|x-y\|_p^2 + \|y-w\|_p^2 + \|w-z\|_p^2 + \|z-x\|_p^2$*

Please refer to [2] for the proof of Lemma 2.

Lemma 3 *Let A_i denote the set of anti-edges at level i and let $\{s,t\} = V(G_0)$, then for $1 < p \leq 2$ and for any $f : G_k \to L_p$,*

$$\|f(s)-f(t)\|_p^2+(p-1)\sum_{i=1}^{k}\sum_{(x,y)\in A_i}\|f(x)-f(y)\|_p^2 \leq \sum_{(x,y)\in E(G_k)}\|f(x)-f(y)\|_p^2$$

.

Sketch of Proof: Let (a,b) be an edge of level i and (c,d) its corresponding anti-edge. By Lemma 2, $\|f(a)-f(b)\|_p^2 + (p-1)\|f(c)-f(d)\|_p^2 \leq \|f(a)-f(c)\|_p^2 + \|f(b)-f(c)\|_p^2 + \|f(d)-f(a)\|_p^2 + \|f(d)-f(b)\|_p^2$. Summing over all such edges and all $i = 0, \cdots, k-1$ yields the desired result by noting that the terms $\|f(x)-f(y)\|_p^2$ corresponding to $(x,y) \in E(G_i)$ cancel for $i = 0, \cdots, k-1$. ■

Theorem 4 *For any $1 < p \leq 2$, any embedding of G_k into L_p incurs distortion at least sqrt$1+(p-1)k$.*

Sketch of Proof: Let $f : G_k \to L_p$ be a non-expansive D-embedding. Since $|A_i| = 4^{i-1}$ and the length of a level i anti-edge is 2^{1-i}, apply Lemma 3 yields $\frac{1+(p-1)k}{D^2} \leq 1$. ■

2 An elementary proof of the Johnson-Lindenstrauss Lemma

Theorem 5 ***(Johnson-Lindenstrauss lemma)*** *For any $0 < \epsilon < 1$ and any integer n, let k be a positive integer s.t.*

$k > 4(\epsilon^2/2 - \epsilon^3/3)^{-1} \ln n$.

Then for any set V of n points in R^d, there is a map $f : R^d \to R^k$ s.t. for all $u, v \in V$,

$(1-\epsilon)\|u-v\|^2 \leq \|f(u) - f(v)\|^2 \leq (1+\epsilon)\|u-v\|^2$.

Further this map can be found in randomized polynomial time.

Let $X_1, \cdots, X_d$ be d independent $N(0,1)$ random variables, and let $Y = \frac{1}{\|x\|}(X_1, \cdots, X_d)$. Let vector $Z \in R^k$ be the projection of Y onto its first k coordinates, and let $L = \|Z\|^2$. Clearly the expected length $\mu = EL = k/d$. Then, we have the following lemma.

Lemma 6 *Let $k < d$. Then*

(a). If $\beta < 1$ then

$Pr[L \leq \beta k/d] \leq \beta^{k/2}(1 + \frac{(1-\beta)k}{(d-k)})^{(d-k)/2} \leq exp(\frac{k}{2}(1-\beta+\ln\beta))$.

(a). If $\beta > 1$ then

$Pr[L \geq \beta k/d] \leq \beta^{k/2}(1 + \frac{(1-\beta)k}{(d-k)})^{(d-k)/2} \leq exp(\frac{k}{2}(1-\beta+\ln\beta))$.

Details of the proof of Lemma 6 can be found in [1].

Sketch of Proof: (Theorem 5)

Only need to consider the case $d > k$. Take a random k-dimensional subspace S, and let v_i' be the projection of vertex $v_i \in V$ into S. Then setting $L = \|v_i' - v_j'\|$ and $\mu = (k/d)\|v_i' - v_j'\|^2$, and applying Lemma 6 we have $Pr[L \leq (1-\epsilon)\mu] \leq 1/n^2$ and $Pr[L \geq (1+\epsilon)\mu] \leq 1/n^2$.

Now choose the map $f(v_i) = (\sqrt{(n/k)}v_i'$. So for fixed pair i, j, the distortion $\|f(v_i) - f(v_j)\|^2 / \|v_i - v_j\|^2$ does not lie in the range $[(1-\epsilon, 1+\epsilon)]$ is at most $1/n^2$. Using the trivial union bound, the chance that some pair of vertices suffers a large distortion is at most $\begin{pmatrix} n \\ 2 \end{pmatrix} \times 2/n^2 = (1 - \frac{1}{n})$. Hence f has the desired properties with probability at least $1/n$. Repeating this projection $O(n)$ times can boost the success probability to any desired constant and give us the claimed randomized polynomial time algorithm.

■

References

[1] Dasgupta, S. and Gupta, A., "An elementary proof of the Johnson-Lindenstrauss Lemma," *http://citeseer.ist.psu.edu/dasgupta99elementary.html.*

[2] Lee, J. and Naor, A.,AuthorName, "Embedding the diamond graph in L_p and dimension reduction in L_1," *Manuscript, June 2003, http://citeseer.ist.psu.edu/lee03embedding.html.*