Approx Nearest Neighbors in High Dims in L_∞

$\left(d_\infty(p,q) = \max_{i=1}^{d} |p_i - q_i| \right)$

LSH or dim reduction don't work well for L_∞

Indyk '98:

space $\tilde{O}(n^t)$ for any $t > 1$. ($\tilde{O}$ hides $d^{O(1)}$ factors)

query time $\tilde{O}(\log n)$

approx factor $4 \lceil \log_t \log(4d) \rceil + 1$

e.g. space $\tilde{O}(n^{1+\delta})$ for any const $\delta > 0$

approx $O(\log \log d)$

or space $\tilde{O}(n^{\log d})$ [note: lower bd by Andoni, Croitoru, Patrascu '08]

approx 5 (can be improved to 3)

idea - back to geom. search trees!

Suffice to solve approx decis problem for fixed r

w.l.o.g., say $r = 1$

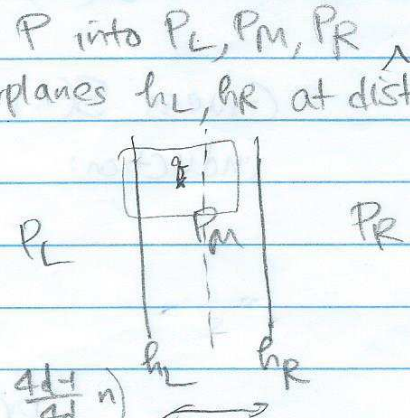
Preproc(P):

Option 1: find good division of P into P_L, P_M, P_R s.t. $|P_L|, |P_R| \geq \frac{n}{4d}$.

by 2 parallel hyperplanes h_L, h_R at dist 2.

Preproc($P_L \cup P_M$)

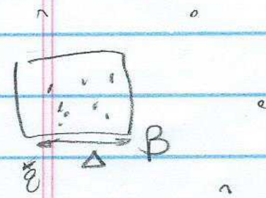
Preproc($P_R \cup P_M$)



$\left(|P_L \cup P_M|, |P_R \cup P_M| \leq \frac{4d-1}{4d} n \right)$

Option 2: find good hypercube B of side length $\le \Delta$ s.t. $|P \cap B| \ge n/2$.

Preproc $(P - B)$



Query(P, q):

Option 1: if q left of $\frac{h_L + h_R}{2}$

Query$(P_L \cup P_M, q)$

else Query$(P_R \cup P_M, q)$

Option 2: if $d(q, B) \le 1$

return any pt in B.

else Query$(P - B, q)$

approx factor $\le \Delta + 1$.

query time $\tilde{O}(\text{tree height}) = \tilde{O}\left(\log_{\frac{4d}{4d-1}} n\right) = \tilde{O}(d \log n)$

Space: let $|P_L| = \alpha n$, $|P_R| = \beta n$, $|P_M| = \delta n$.

$$\Rightarrow S(n) = \max \begin{cases} S((\alpha+\delta)n) + S((\beta+\delta)n) \cancel{+1} \\ S(n/2) + 1 \end{cases}$$

Guess $S(n) \le n^t$

Induction: $S(n) \le (\alpha+\delta)^t n^t + (\beta+\delta)^t n^t$

$= \left((1-\beta)^t + (\beta+\delta)^t\right) n^t$

$\le \left(1 - \beta + (\beta+\delta)^t\right) n^t$

$\le n^t$

provided $\boxed{\beta \ge (\beta+\delta)^t}$

or similarly, $\boxed{\alpha \ge (\alpha+\delta)^t}$.

<u>Claim</u> For any pt set P,

(1) $\exists$ good division with $\alpha \geq (\alpha+\delta)^t$ or $\beta \geq (\beta+\delta)^t$

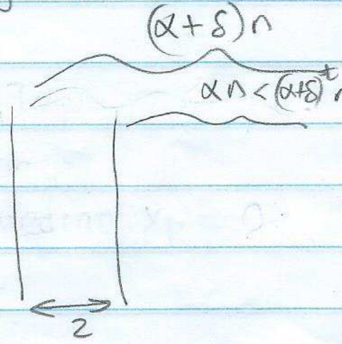
OR (2) $\exists$ good hypercube B of side length

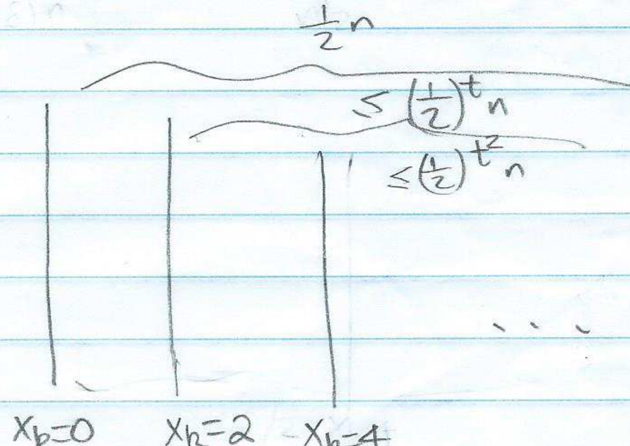
$$\Delta = 4 \lceil \log_t \log(4d) \rceil.$$

Pf: Assume (1) false.

Fix $k \in [d]$.

By translation, assume median x_k is 0.





$$\Rightarrow \left(\#\text{ pts with } x_k \geq 2\ell\right) \leq \left(\frac{1}{2}\right)^{t^\ell} n \leq \frac{n}{4d}$$

by setting $\ell = \lceil \log_t \log(4d) \rceil$.

Similarly, $\left(\#\text{ pts with } x_k \leq -2\ell\right) \leq \frac{n}{4d}$

Choose $B = [-2\ell, 2\ell]^d$.

$$\Rightarrow \quad |P - B| \leq d\left(\frac{n}{4d} + \frac{n}{4d}\right) = \frac{n}{2} \qquad \square$$