Computational Geometry Lecture 37

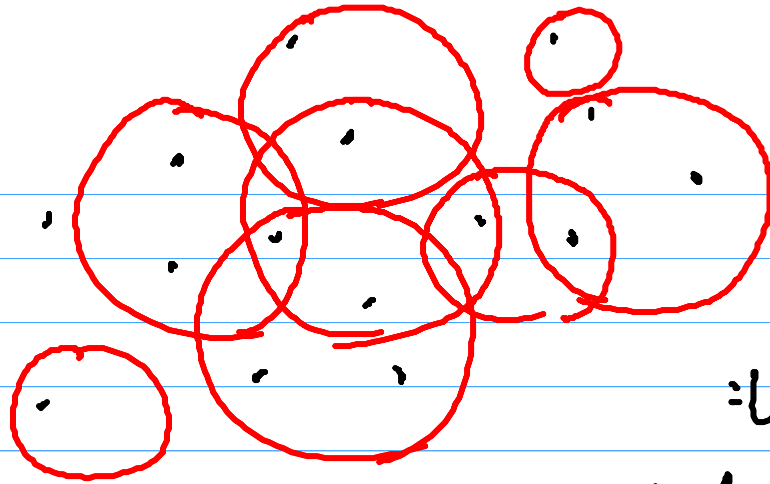
Topic : Geometric set cover

$S = (X, R)$ has S has finite VC dimension.

X : set of n points
R : " " m ranges

We want to choose a subset $R' \subset R$ that covers all points of X, namely $\forall x \in X$, $\exists r \in R'$, $x \in r$

We can use the general set cover algorithm (say the greedy algorithm) to find a cover of size $k \cdot \log n$ $\quad k = |\text{OPT cover}|$



This problem is known to be NP-hard.: finding the smallest size set cover.

Can we improve upon the generic set cover soln?

$$S = (X, R)$$

Defn: <u>ε-net</u>. A subset $N \subseteq X$ is a ε-net for X, if for any range $r \in R$ and $m(r) \geq \varepsilon$

then N contains at least one point of $r \cap X$.

$$m(r) = \frac{r \cap X}{|X|}$$

$m(r) \geq \varepsilon$

$$\left| \frac{|N \cap r|}{|N|} - \frac{|X \cap r|}{|X|} \right| \leq \phi$$

How does one construct an ε-net?

<u>Theorem</u> Let (X, R) be a range space of VC dimension δ and x be a finite subset of X.
For $0 < \varepsilon < 1$ and $\varphi < 1$ let N be a subset obtained by m independent draws from x (uniformly chosen), where

$$m \geq \max\left\{\frac{4}{\varepsilon}\log\left(\frac{4}{\varphi}\right), \frac{8\delta}{\varepsilon}\log\left(\frac{16}{\varepsilon}\right)\right\}$$

then N is an ε-net with probability $\geq 1-\varphi$

$|N| = m$

Suppose we use the shattering dimension of S, say d.
Then sample size $\geq O\left(\frac{d}{\varepsilon}\log\frac{d}{\varepsilon}\right)$

<u>Comment:</u> Suppose the elements of $\mathcal{X}$ are weighted, i.e. there is a weight function

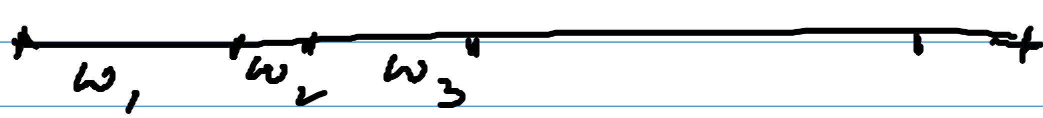
$$w : \mathcal{X} \to \mathbb{R}^+ . \quad w(r) = \sum_{y \in r} w(y)$$

Suppose

$$W = \sum_{y \in \mathcal{X}} w(y)$$

then we may be interested to hit the "large weighted" ranges, namely those whose weights are $\geq \varepsilon \cdot W$

A similar result can be proved where the ordinary sampling is replaced with "weighted sampling"



Algorithm Set cover

It repeatedly selects an ε-net for some "ε"

$\exists$ S^* : suppose has a shattering dimension δ^*

- Choose an ε-net (weighted) of size $O\left(\frac{\delta^*}{\varepsilon}\log\left(\frac{\delta^*}{\varepsilon}\right)\right)$, call it Y

- Verify if Y is an ε-net : if not discard

: else (if Y is an ε-net), then we check if Y is a cover (brute force)

If it is a cover, then output $\underline{\underline{Y}}$

else (if not a cover)

then there must be at least one point $p \in X$ that is not covered by Y.

let $R_p = \{ r \in R \mid p \in r \}$

[Note that R_p must include same range $r' \in OPT$

We will double the weight of elements in R_p if $w(R_p) < \varepsilon W$.

Observation: Every time, we double the weight, we are increasing the weight by no more than $(1+\varepsilon)$ multiplicative factor

Therefore $W_i \leq (1+\varepsilon) W_{i-1}$

$$\Rightarrow W_i \leq (1+\varepsilon)^i m = W_0$$

(i is an iteration where weights are doubled)