Algorithmic Excursions: Topics in Computer Science II **Spring 2016**

Lecture 3 & 4 : ε-net(contd.), ε-approximation and Discrepancy

Lecturer: Kasturi Varadarajan *Scribe: Santanu Bhowmick*

In the last lecture, we looked at a probabilistic proof of the following lemma, for which we now provide a deterministic algorithm.

Lemma 3.1 *Let $\mathcal{S} = (X, \mathcal{R})$ be a finite range space and $0 < \epsilon < 1$. Then $\mathcal{S}$ has an ϵ-net of size $O(\frac{1}{\epsilon}\ln|\mathcal{R}|)$.*

Proof: We construct a set $N \subseteq X$ using the following algorithm.

$R' \leftarrow \{r \in R \mid |r| > \varepsilon \cdot |X|\}$.
$N \leftarrow \varnothing$.
while $R' \neq \varnothing$ **do**
    Pick $x \in X$ that occurs in maximum number of ranges in $\mathcal{R}$.
    $R' \leftarrow R' \setminus \{r \in R' \mid x \in r\}$.
    $N \leftarrow N \cup \{x\}$
return N

By construction, N is an ε-net as it has at least one element from each "sufficiently large" range i.e. ranges having more than $\varepsilon \cdot |X|$ elements. We bound the size of N as follows.

Suppose $|R'| = k$ at the beginning of a certain iteration. Each $r \in R'$ contains more than $\varepsilon \cdot |X|$ elements by definition. We claim that there exists an element $x \in X$ that is contained in at least $\frac{\varepsilon \cdot |X|}{|X|} \cdot k$ ranges in R'. (To see why, consider the directed bipartite graph $G = (R', X, E)$ where $(r, x) \in E$ if $x \in r$. Each vertex $r \in R'$ has at least $\varepsilon \cdot |X|$ outgoing edges, so the total outdegree from the set R' is at least $\varepsilon \cdot |X| \cdot k$. The average indegree of X is thus $\frac{\varepsilon \cdot |X|}{|X|} \cdot k = \varepsilon \cdot k$. Hence, there exists at least one element in X with indegree $\varepsilon \cdot k$.) Thus, after the iteration of the while loop,

$$|R'| \leq k - \varepsilon \cdot k = (1 - \varepsilon) \cdot k$$

Homework: Show that the bound on the size of N in the lemma follows. ■

Lemma 3.2 *Suppose $\mathcal{S} = (X, \mathcal{R})$ has VC-dimension d, and let $\Pi_{\mathcal{S}}$ be its shatter function. Then*

$$\Pi_{\mathcal{S}}(m) \leq \binom{m}{0} + \binom{m}{1} + \binom{m}{2} + \cdots + \binom{m}{d} \equiv \phi_d(m)$$

Proof: Let $Y \subseteq X$ be a finite subset with m elements. Then $(Y, \mathcal{R}_Y)$ has VC-dimension at most d. It suffices to show that if $(X, \mathcal{R})$ is a finite range space with VC-dimension at most d and $|X| = m$, then $|\mathcal{R}| \leq \phi_d(m)$. We prove the result by induction on m and d as follows.

Fix $x \in X$, and let $\mathcal{R}_1 = \mathcal{R}_{X \setminus \{x\}}$. The range space $(X \setminus \{x\}, \mathcal{R}_1)$ has VC-dimension of at most d. So, inductively, $|\mathcal{R}_1| \leq \phi_d(m-1)$.

Let $\mathcal{R}_2 = \{r \in R \mid x \notin r, r \cup \{x\} \in \mathcal{R}\}$. The VC-dimension of $(X \setminus \{x\}, \mathcal{R}_2)$ is at most $d - 1$, and hence $|\mathcal{R}_2| \leq \phi_{d-1}(m-1)$.

Then,

$$
\begin{aligned}
|\mathcal{R}| &= |\mathcal{R}_1| + |\mathcal{R}_2| \\
&\le \phi_d(m-1) + \phi_{d-1}(m-1) \\
&= \phi_d(m)
\end{aligned}
$$

The last equality can be explained by the component-wise sum of the two terms $\phi_d(m-1)$ and $\phi_{d-1}(m-1)$, as follows:

$$
\begin{array}{rcccccccc}
\phi_d(m-1) = & \binom{m-1}{0} + \binom{m-1}{1} & + & \binom{m-1}{2} & + & \ldots & + \binom{m-1}{d-1} & + \binom{m-1}{d} \\
\phi_{d-1}(m-1) = & \binom{m-1}{0} & + & \binom{m-1}{1} & + & \ldots & + \binom{m-1}{d-2} & + \binom{m-1}{d-1} \\
\hline
\phi_d(m) = & \binom{m}{1} & + & \binom{m}{2} & + & \ldots & + \binom{m}{d-1} & + \binom{m}{d}
\end{array}
$$

■

It follows that if $|\mathcal{R}| = O(|X|^d)$, then ε-net size is $O(\frac{d}{\varepsilon} \log |X|)$.

Definition 3.3 *Let $\mathcal{S} = (X, \mathcal{R})$ be a finite range space, and let $0 \le \varepsilon \le 1$. A subset $A \subseteq X$ is an **ε-sample** (**ε-approximation**) if for any $r \in R$,*

$$
\left| \frac{|X \cap r|}{|X|} - \frac{|A \cap r|}{|A|} \right| \le \varepsilon
$$

Lemma 3.4 *If A is an ε-approximation, it is also an ε-net.*

Proof: Let $r \in \mathcal{R}$ be a range having greater than $\varepsilon \cdot |X|$ elements. Then, $\frac{|X \cap r|}{|X|} > \varepsilon$, and since A is an ε-approximation, $\frac{|A \cap r|}{|A|} > 0$ and thus A has non-zero intersection with range r. ■

Definition 3.5 *Let $\mathcal{S} = (X, \mathcal{R})$ be a range space. Let $\chi : X \to \{-1, +1\}$ be a coloring. We denote/define/say:*

- *For $r \in \mathcal{R}$, let $\chi(r) \equiv \sum_{x \in r} \chi(x)$.*
- *Discrepancy of χ over $r \equiv |\chi(r)|$.*
- *Discrepancy of χ, disc$(\chi) \equiv \max_{r \in R} |\chi(r)|$.*
- *Discrepancy of $\mathcal{S} \equiv \min_{\chi : X \to \{-1,+1\}}$ disc(χ).*

Definition 3.6 *Suppose $|X|$ is even, and Π is a partition of X into pairs. We can say that $\chi : X \to \{-1, +1\}$ is compatible with Π if for each $\{p, q\} \in \Pi$, either*

- *$\chi(p) = +1$ and $\chi(q) = -1$, or*
- *$\chi(q) = +1$ and $\chi(p) = -1$*

Lemma 3.7 *Let $\mathcal{S} = (X, \mathcal{R})$ be a range space, and let Π be a partition of X into pairs. Let $|X| = n, |\mathcal{R}| = m$. Let χ be a random coloring compatible with Π. Then* $\mathbf{Pr}\left[disc(\chi) < \sqrt{n \cdot \ln 4m}\right] \geq \frac{1}{2}$.

Proof: For range $r \in \mathcal{R}$, let $\{x_1, x_2, \ldots, x_t\} \subseteq r$ be those elements paired with an element not in r. Then

$$\chi(r) = \chi(x_1) + \chi(x_2) + \cdots + \chi(x_t)$$

is the sum of t independent random variables uniformly chosen from $\{-1, +1\}$. Hence, for any $\Delta > 0$, we have the following by applying Chernoff bound,

$$\mathbf{Pr}\left[\chi(r) \geq \Delta\right] < e^{\frac{-\Delta^2}{2t}} = \frac{1}{e^{\frac{\Delta^2}{2t}}}$$

Setting $\Delta = \sqrt{2t \cdot \ln 4m}$, we get $\mathbf{Pr}[\chi(r) \geq \sqrt{2t \cdot \ln 4m}] < \frac{1}{e^{\ln 4m}} = \frac{1}{4m}$.

Since $t \leq \frac{n}{2}$,

$$\mathbf{Pr}\left[\chi(r) \geq \sqrt{n \cdot \ln 4m}\right] \leq \mathbf{Pr}\left[\chi(r) \geq \sqrt{2t \cdot \ln 4m}\right] < \frac{1}{4m}$$

By symmetry, we get

$$\mathbf{Pr}\left[|\chi(r)| \geq \sqrt{n \cdot \ln 4m}\right] \leq \frac{2}{4m} = \frac{1}{2m}.$$

Finally, using Union bound, we have

$$\begin{aligned}
\mathbf{Pr}\left[\text{disc}(\chi) \geq \sqrt{n \cdot \ln 4m}\right] &= \mathbf{Pr}\left[\bigcup_r |\chi(r)| \geq \sqrt{n \cdot \ln 4m}\right] \\
&\leq \sum_r \mathbf{Pr}\left[|\chi(r)| \geq \sqrt{n \cdot \ln 4m}\right] \\
&\leq \sum_r \frac{1}{2m} = \frac{1}{2}
\end{aligned}$$

from which the claim follows. ■

Notes on concentration measures: We explore the difference between using Chernoff bound and using Chebyshev's Inequality in this short example.

Let $Y = Y_1 + Y_2 + \cdots + Y_t$, where each Y_i is chosen independently uniformly at random from $\{-1, +1\}$. By Chernoff bound, $\mathbf{Pr}[Y \geq \Delta] < \frac{1}{e^{\Delta^2/2t}}$.

We now bound the same probability using Chebyshev's inequality. We note that $\mathbf{E}[Y_i] = 0, \mathbf{E}[Y] = \sum_i \mathbf{E}[Y_i] = 0, \mathbf{E}[Y_i^2] = 1, \mathbf{Var}[Y_i] = \mathbf{E}[Y_i^2] - (\mathbf{E}[Y_i])^2 = 1$. Due to independence of each Y_i, $\mathbf{Var}(Y) = \sum_{i=1}^{t} \mathbf{Var}(Y_i) = t$, and $\sigma(Y) = \sqrt{\mathbf{Var}(Y)} = \sqrt{t}$.

By Chebyshev Inequality, for any real number $\alpha > 0$, $\mathbf{Pr}\left[|Y - \mathbf{E}[Y]| \geq \alpha \cdot \sqrt{t}\right] \leq \frac{1}{\alpha^2}$.

Plugging $\alpha = 10$ in the above, we get

$$\mathbf{Pr}[|Y| \geq 10 \cdot \sqrt{t}] \leq \frac{1}{100}$$

Plugging $\Delta = 10 \cdot \sqrt{t}$ in the inequality obtained using Chernoff bound, we have

$$\mathbf{Pr}[Y \geq 10 \cdot \sqrt{t}] < \frac{1}{e^{50}} \Rightarrow \mathbf{Pr}[|Y| \geq 10 \cdot \sqrt{t}] < \frac{2}{e^{50}}$$

Thus, the bound obtained using Chernoff bound is a much more precise bound than that obtained using Chebyshev Inequality.

Lemma 3.8 *Given a range space $\mathcal{S} = (X, \mathcal{R})$, a partition Π, a coloring χ compatible with Π and disc$(\chi) \leq f$, let $Q = \{x \in X \mid \chi(x) = -1\}$. Then, for any $r \in R$, $\left|\frac{|X \cap r|}{|X|} - \frac{|Q \cap r|}{|Q|}\right| \leq \frac{f}{n}$, i.e. Q is an $\frac{f}{n}$-approximation.*

Proof: Fix $r \in \mathcal{R}$. Then,

$$\begin{aligned} |\chi(r)| &= ||X \setminus Q \cap r| - |Q \cap r|| \\ &= ||X \cap r| - |Q \cap r| - |Q \cap r|| \\ &= ||X \cap r| - 2 \cdot |Q \cap r|| \leq f \end{aligned}$$

Dividing last inequality by $|X| = 2 \cdot |Q| = n$, we get

$$\left|\frac{|X \cap r|}{|X|} - \frac{2 \cdot |Q \cap r|}{2 \cdot |Q|}\right| \leq \frac{f}{n}$$

■

Lemma 3.9 *If A is an ε-approximation for $(X, \mathcal{R})$ and A' is an ε'-approximation for $(A, \mathcal{R}_A)$, then A' is an $(\varepsilon + \varepsilon')$-approximation for $(X, \mathcal{R})$.*

The proof of the above claim is left as a **homework**.