The ε-net Theorem

1 Hitting Set Wrap-Up

Recall that using the multiplicative weights update algorithm, we were able to find a weight vector $\boldsymbol{w}^*$ that gives us a hitting set when we know OPT. However, this is not too much of a problem: since $OPT \leq n$, we can use a doubling search where we check $OPT = 1, 2, 4, \ldots$. Once it is too large, we can use binary search to find a tighter bound. To diagnose whether OPT is too small, we can observe whether multiplicative weights update (with $\varepsilon = OPT/2$) terminates in $4n$ rounds. If not, then ε was too large, which means OPT was too small. If so, our guess of OPT is within a factor 2 of the actual OPT, so we can binary search (or just use that guess). Either way, we obtain an $\Theta(OPT \cdot \log OPT)$ approximation.

If the set system $(X, \mathcal{C})$ has the property that an ε-net of smaller size than $\mathcal{O}(\frac{d}{\varepsilon} \log \frac{d}{\varepsilon})$ can be found efficiently – for instance, $\mathcal{O}(\frac{d}{\varepsilon})$ – then the algorithm's performance improves to a constant-factor approximation.

2 The ε-net Theorem

Theorem. *Let (Ω, R) be a set system with $n = |\Omega|$ and VC-dimension d. Let $N \subseteq \Omega$ be an i.i.d. random sample of Ω of size*

$$m = |N| = \frac{8d}{\varepsilon} \log \frac{8d}{\varepsilon} + \frac{4}{\varepsilon} \log \frac{2}{\delta}.$$

Then, N is an ε-net with probability at least $1 - \delta$.

Proof. We will show the proof only for the uniform distribution over Ω. Define

$$g(d, n) := \sum_{k=0}^{d} \binom{n}{k}.$$

We will use the following lemmas:

Lemma 1. *[Sauer-Shelah]. If (Ω, R) has VC-dimension d, then $|R| \leq g(d, n)$.*

Proof. Nontrivial by induction on n using the effect of removing one x from Ω. □

Lemma 2. *$g(d, n) \leq n^d$ for all n and d.*

Proof. Induction and simple bounds on binomial coefficients. □

Recall also that for any set system (Ω, R) with $X \subseteq R$, the projection of (Ω, R) onto X is defined as

$$R|_X = \{S \cap X : S \in R\}.$$

Lemma 3. *The VC-dimension of $(X, R|_X)$ is at most the VC-dimension of (Ω, R).*

Proof. Obvious. □

We want to bound (by δ) the probability of the event that we do not obtain an ε-net because some large set is not hit:

$$\mathcal{E} := [\exists A \in R : |A| \geq \varepsilon n, A \cap N = \emptyset].$$

Consider sampling a second i.i.d. set T of the same size as N (both are multisets, so we keep duplicates). Define the event that there is a large set that N does not hit, but T samples well:

$$\mathcal{F} := [\exists A \in R : |A| \geq \varepsilon n, A \cap N = \emptyset, \text{ and } |A \cap T| \geq \frac{\varepsilon m}{2}].$$

Since the expected $|A \cap T|$ is εm, this should be pretty likely.

Even though $\mathcal{F}$ looks more difficult, it is actually easier to analyze, so we will bound $\Pr[\mathcal{F}]$ instead of $\Pr[\mathcal{E}]$.

Lemma 4. $\Pr[\mathcal{F}] \geq \frac{1}{2}\Pr[\mathcal{E}]$.

Proof. Because $\mathcal{F} \subseteq \mathcal{E}$, we have $\Pr[\mathcal{F}] \leq \Pr[\mathcal{E}]$ and $\Pr[\mathcal{F} \mid \mathcal{E}] = \Pr[\mathcal{F}]/\Pr[\mathcal{E}]$. So $\Pr[\mathcal{F}] = \Pr[\mathcal{E}] \cdot \Pr[\mathcal{F}|\mathcal{E}]$. We will show that $\Pr[\mathcal{F}|\mathcal{E}] \geq 1/2$.

Because we conditioned on $\mathcal{E}$, there exists an A with $|A| = \varepsilon n$ and $N \cap A = \emptyset$. Fix one such A and show that with probability at least $1/2$ we have

$$|A \cap T| \geq \frac{\varepsilon m}{2}.$$

We will use tail bounds to show this. Write

$$|A \cap T| = \sum_{i=1}^{m} z_i$$

where z_i is the indicator random variable for whether $i \in A$. We have

$$\begin{aligned}
\mathbb{E}[|A \cap T|] &= \sum_{i=1}^{m} \Pr[i \in A] \\
&= m \Pr[A] \\
&= \varepsilon m.
\end{aligned}$$

We will use Chebyshev, so we need to get the variance of $|A \cap T|$. Because we have pairwise independence and we assume $\varepsilon \leq 1/2$,

$$\begin{aligned}
Var[|A \cap T|] &= \sum_{i=1}^{m} m \Pr[A](1 - \Pr[A]) \\
&= m\varepsilon(1 - \varepsilon).
\end{aligned}$$

So by Chebyshev,

$$\begin{aligned}
\Pr[|A \cap T|] < \frac{\varepsilon m}{2}] &\leq \Pr[||A \cap T| - \mathbb{E}[|A \cap T|]| > \frac{\varepsilon m}{2}] \\
&\leq \frac{Var[|A \cap T|]}{(\frac{\varepsilon m}{2})^2} \\
&= \frac{4m\varepsilon(1-\varepsilon)}{\varepsilon^2 m^2} \\
&= \frac{4(1-\varepsilon)}{\varepsilon m} \\
&\leq \frac{4(1-\varepsilon)}{\varepsilon \frac{8d}{\varepsilon}} \\
&= \frac{1}{2d} \\
&\leq \frac{1}{2}.
\end{aligned}$$

Thus, $\Pr[\mathcal{F} \mid \mathcal{E}] \geq 1/2$. ∎

We will now bound the probability of $\mathcal{F}$. Then, $\Pr[\mathcal{E}] \leq 2\Pr[\mathcal{F}]$. Recall that $\mathcal{F}$ is the event that there is a set A completely missed by the first m random samples N, but well-sampled by the second m random samples T.

The idea is to first sample $Z := N \cup T$ (a set of $2m$ elements), then partition them uniformly randomly into N and T. We are only partitioning constantly many elements, so we can apply Sauer-Shelah and the union bound.

For every set $A \in R$ with $|A| \geq \varepsilon n$, define the event that A is a witness:

$$\mathcal{F}_A = [A \cap N = \emptyset \text{ and } |A \cap T| \geq \frac{\varepsilon m}{2}].$$

Thus

$$\mathcal{F} = \bigcup_{A:|A|\geq \varepsilon n} \mathcal{F}_A.$$

We want to take a union bound to bound $\Pr[\mathcal{F}]$.

If $|A \cap Z| < \frac{\varepsilon m}{2}$, then $\mathcal{F}_A$ cannot happen because then $|A \cap T| < \frac{\varepsilon m}{2}$. So we only need to focus on sets A with $|A \cap Z| \geq \frac{\varepsilon m}{2}$. These are sets that are large with respect to the new ground set Z. For any such A, the event $\mathcal{F}_A$ requires us to put all elements from $A \cap Z$ into T and none into N. Thus,

$$\begin{aligned}
\Pr[\mathcal{F}_A] &\leq \frac{\binom{2m-\frac{\varepsilon m}{2}}{m}}{\binom{2m}{m}} \\
&= \frac{\frac{(2m-\frac{\varepsilon m}{2})!}{m!(m-\frac{\varepsilon m}{2})!}}{\frac{(2m)!}{m!\cdot m!}} \\
&= \frac{(2m-\frac{\varepsilon m}{2})\dots(m+1)(m)\dots(m-\frac{\varepsilon m}{2}+1)}{2m(2m-1)\dots(2m-\frac{\varepsilon m}{2}+1)(2m-\frac{\varepsilon m}{2})\dots(m+1)} \\
&= \frac{m(m-1)\dots(m-\frac{\varepsilon m}{2}+1)}{2m(2m-1)\dots(2m-\frac{\varepsilon m}{2}+1)} \\
&\leq 2^{-\frac{\varepsilon m}{2}}.
\end{aligned}$$

Consider A, A' with $|A|, |A'| \geq \varepsilon n$. If $A \cap Z = A' \cap Z$, then conditioned on Z we have $\mathcal{F}_A = F_{A'}$. So, our union bound only needs to go over all distinct sets $A \cap Z$ (not all A). In other words, over at most $|R|_Z|$ sets.

By the earlier lemmas, $(Z, R|_Z)$ has VC-dimension at most d, so by Sauer-Shelah,

$$\begin{aligned}
|R|_Z| &\leq g(d, 2m) \\
&\leq (2m)^d.
\end{aligned}$$

So $\Pr[\mathcal{F}] \leq (2m)^d \cdot 2^{-\frac{\varepsilon m}{2}}$. It remains to show that $(2m)^d \cdot 2^{-\frac{\varepsilon m}{2}} \leq \delta/2$ (because $\Pr[\mathcal{E}] \leq 2\Pr[\mathcal{F}]$). Taking logs,

$$\frac{\varepsilon m}{2} \geq \log\frac{2}{\delta} + d\log(2m).$$

Recall

$$\begin{aligned}
m &= \frac{8d}{\varepsilon}\log\frac{8d}{\varepsilon} + \frac{4}{\varepsilon}\log\frac{2}{\delta} \\
&\geq \frac{4}{\varepsilon}\log\frac{2}{\delta},
\end{aligned}$$

so

$$\frac{\varepsilon m}{4} \geq \log\frac{2}{\delta}.$$

Then we need to show

$$\frac{\varepsilon m}{4} \geq d \log(2m).$$

Since the left-hand side grows faster than the right-hand side with respect to m, we just need to find some m_0 where this holds. With

$$m_0 = \frac{8d}{\varepsilon} \log \frac{8d}{\varepsilon},$$

we get

$$\frac{\varepsilon m_0}{4} = 2d \log \frac{8d}{\varepsilon}$$

and

$$\begin{aligned} d \log(2m_0) &= d \log(\frac{16d}{\varepsilon} \log \frac{8d}{\varepsilon}) \\ &= d(1 + \log \frac{8d}{\varepsilon} + \log \log \frac{8d}{\varepsilon}). \end{aligned}$$

Since $1 + \log \log \frac{8d}{\varepsilon} \leq \log \frac{8d}{\varepsilon}$, the inequality holds. Thus, we have shown

$$\Pr[\mathcal{F}] \leq \frac{\delta}{2}.$$

$\square$