Convergence Based on VC Dimension

These notes are slightly edited from previous scribe notes (in Spring 2006) taken by Alexander Ecker.

1 $\epsilon$-Nets

In this lecture, we proved the following Theorem:

1.1 Theorem

If a sample of size $m \geq \max \left\{ \frac{6}{\epsilon} \ln \frac{2}{\delta}, \frac{3d}{\epsilon} \ln \frac{30}{\delta} \right\}$ is drawn iid. via $D$, then, with probability at least $1 - \delta$, $S$ is an $\epsilon$-Net.

Various definitions as well as a high level outline of the proof is given in the slides for the lecture. In these notes we only give the proofs of the five claims used in the proof.

1.2 Claim 1

If $m \geq \frac{6}{\epsilon} \ln 2$, then $P[A] \leq 2P[B]$.

Proof: We observe that $A \subset B$, hence $P[B] = P[A,B]$. The conditional probability $P[B|A]$ can be written as

$$P[B|A] = P \left[ \text{m examples with } p = P(\text{success}) \geq \epsilon, \#(\text{successes}) \geq \frac{em}{2} \right]$$

$$= 1 - P \left[ \text{m examples with } p \geq \epsilon, \hat{p} < \frac{\epsilon}{2} \right]$$

$$= 1 - P \left[ \hat{p} < p \left( 1 - \frac{1}{2} \right) \right]$$

$$\geq 1 - e^{-\frac{mc}{2}} > \frac{1}{2}$$

Note that the following variant of the Chernoff Bound was used:

$$P \left[ \hat{p} < p(1 - \gamma) \right] < e^{-\frac{m\gamma^2}{2}}$$

It follows that

$$P[B] = P[A,B] = P[\text{A}] P[\text{B}|\text{A}] \geq P[\text{A}] \cdot \frac{1}{2}$$

$$\Rightarrow P[\text{A}] \leq 2P[\text{B}]$$
1.3 Claim 2

\[ P[B] \leq \max_S P[B'(S)]. \]

**Proof:**

\[
P[B] = \sum_S P[S \text{ sampled}] P[B|S \text{ sampled}] \leq \max_S P[B'(S)] \sum_{S \text{ sampled}} P[S \text{ sampled}] = P[B'(S)]
\]

\[= \]

1.4 Claim 3

\[ P[B'(S)] \leq |\Pi_{\Delta_c(f)}(S)| \cdot \max_{\hat{r}} P[\hat{r} \subseteq S_2] \text{ where } \hat{r} \in \Pi_{\Delta_c(f)}(S) \text{ and } |\hat{r}| \geq \frac{cm}{2}. \]

**Proof:**

\[
P[B'(S)] = \sum_{\hat{r} \in \Pi_{\Delta_c(f)}(S) \text{ and } |\hat{r}| \geq \frac{cm}{2}} \sum_{\hat{r} \subseteq S_2} P[\hat{r} \subseteq S_2] \leq (\#\hat{r}) \cdot \max_{\hat{r}} P[\hat{r} \subseteq S_2]
\]

1.5 Claim 4

For any \( \hat{r} \in \Pi_{\Delta_c(f)}(S) \) such that \( |\hat{r}| \geq \frac{cm}{2} \): \( P[\hat{r} \subseteq S_2] \leq \left( \frac{1}{2} \right)^{\frac{cm}{2}} \).

**Proof:** Fix \( \hat{r} \) such that \( |\hat{r}| = l > \frac{cm}{2} \). First consider all permutations but focusing on locations of \( \hat{r} \).

- Choose locations: \( \binom{2m}{l} \)
- Order locations: \( l! \)
- Place rest: \( (2m - l)! \)

So we can view the number of permutations as \( \Rightarrow (2m)! = \binom{2m}{l}! (2m - l)! \)

Doing the same but restricting permutations so that \( \hat{r} \subseteq S_2 \) we get \( \binom{m}{l}! (2m - l)! \). Now since we are choosing a permutation at random we get

\[
P[\hat{r} \subseteq S_2] \leq \frac{\binom{m}{l}}{\binom{2m}{l}} = \frac{m(m - 1) \cdots (m - l + 1)}{2m(2m - 1) \cdots (2m - l + 1)} \leq \left( \frac{1}{2} \right)^l \leq \left( \frac{1}{2} \right)^{\frac{cm}{2}}
\]

1.6 Claim 5

\[ |\Pi_{\Delta_c(f)}(S)| \leq \Phi_d(2m) \leq \left( \frac{2m}{d} \right)^d. \]

**Proof:** We show:

\[
VCD(C) = VCD(\Delta(f)) = d
\]

This is done by giving a bijection from \( \Pi_C(S) \rightarrow \Pi_{\Delta_c(f)}(S) \). Let \( A \in \Pi_C(S) \), \( A = S \cap c \) for some \( c \in C \). Define \( \tilde{A} = A \Delta (f \cap S) = (c \cap S) \Delta (f \cap S) = (c \Delta f) \cap S \Rightarrow \tilde{A} \in \Pi_{\Delta_c(f)}(S) \). It is easy to see that the mapping is reversible and is therefore a bijection. Since the sizes of \( \Pi_C(S) \) and \( \Pi_{\Delta_c(f)}(S) \) are the same so is the size of the largest shattered set.

It follows that \( VCD(\Delta_c(f)) \leq d \) and therefore \( \Pi_{\Delta_c(f)}(2m) \leq \left( \frac{2m}{d} \right)^d \) where the last inequality was discussed in a previous lecture (and proved in the [KS] text).