# Probabilistic graphical models, Assignment 3

Siddharth Bhat (20161105)

March 21st, 2020

## 6.8, Q1:

Monotonicity of VC dimension

Let $\mathcal{H}' \subseteq \mathcal{H}$. Show that $\text{VCdim}(\mathcal{H}') \leq \text{VCdim}(\mathcal{H})$.

### Answer

Recall that the definition of VCdimis is that $\text{VCdim}(\mathcal{H})$ is the maximal size of a set $C \subseteq \mathcal{X}$ which can be *shattered* by $\mathcal{H}$.

Expanding the definition of shattering, we get that the $\text{VCdim}(\mathcal{H})$ is the maximal size of *any* set $C \subseteq X$ such that $\mathcal{H}$ restricted to $C$ is the set of all functions from $C$ to $\{0, 1\}$.

Now, If $C \subseteq \mathcal{X}$ is shattered by $\mathcal{H}' \subseteq \mathcal{H}$, then this means that:

$$|\{f|_C : f \in H'\}| = 2^{|C|}$$

Since $\mathcal{H}' \subseteq \mathcal{H}$, we can replace $\mathcal{H}'$ with $\mathcal{H}$ in the above formula to arrive at:

$$|\{f|_C : f \in H\}| = 2^{|C|}$$

So, clearly, $\text{VCdim}(\mathcal{H}') \leq \text{VCdim}(\mathcal{H})$. However, there might be a set that is *larger* than $C$ that can be shattered by $\mathcal{H}$. This lets us get the strict equality $\text{VCdim}(\mathcal{H}) < \text{VCdim}(\mathcal{H})$ in certain cases — that is, we *cannot* assert that $\text{VCdim}(H) \leq \text{VCdim}(H')$. For example, if we choose $\mathcal{H}' = \emptyset$ where $\mathcal{H}$ is a hypothesis class with $\text{VCdim}(\mathcal{H}) = 1$. Then $\text{VCdim}(\emptyset) = 0 < 1 = \text{VCdim}(\mathcal{H})$.

## 6.8, Q2:

Given a finite domain $\mathcal{X}$, and a finite number $k \leq |\mathcal{X}|$s, find and prove the VC dimension of:

### A. Functions that assign 1 to exactly $k$ elements of $\mathcal{X}$

$\mathcal{H} \equiv \left\{ h \in \{0,1\}^X : \; |\{x : h(x) = 1\}| = k \right\}.$

### Solution

VC dimension is $k/2$.

### B. Functions that assign 1 to at most $k$ elements of $\mathcal{X}$

#### Solution

VC dimension is $k$.

## 6.8, Q3:

Let $\mathcal{X}$ be the boolean hypercube $\{0,1\}^n$. We define parity to be:

$$h_I : \mathcal{X} \to \{0,1\}; h_I((x_1, x_2, \dots, x_n)) \equiv \sum_{i \in I}(x_i) \mod 2.$$

What is the VC dimensions of the set of all parity functions? That is,

$$\mathcal{H}_{parity,n} \equiv \{h_I : I \subseteq \{1, 2, \dots n\}\}$$

#### Solution

Once again, unwrapping the definition, our hypothesis class can compute the sum modulo 2 of *all of the subsets of* $\vec{x} \in \mathcal{X}$. We need to use this to find the *largest* set $C \subseteq \mathcal{X} \equiv \{0,1\}^n$ such that $|\mathcal{H}_C| = 2^{|C|}$.

We can interpret elements $(h \in H)$ as a vector $h_I \in \{0,1\}^n$, where $h_I$ is a vector with 1's at each index $i \in I$, and 0 at other indexes. That is:

$$h_I \in \{0,1\}^n \qquad h_I[i] \equiv \mathbb{1}[i \in I] = \begin{cases} 1 & i \in I \\ 0 & \text{otherwise} \end{cases}$$

We can reinterpret the function $h_I(x)$ as $h^T x$ where we have a vector space over the galois field $GF_2$, where $\oplus$ denotes XOR (recall that addition mod 2 is XOR).

$$h_I(x) = \bigoplus_{i \in I} x_i = \bigoplus_{i=1}^{n} \mathbb{1}[i \in I]x_i = \bigoplus_{i=1}^{n} h_I[i]x[i] = h_I^T x$$

Now, we can reinterpret the question of finding the VC dimension as finding the largest collection of vectors $C \subseteq \mathcal{X} = \{0,1\}^n$ such that the function $C_{act}$ has full image, where the function $C_{act}$ is:

$$\begin{aligned}
&C_{act} : \mathcal{H} \to \{0,1\}^{|C|} \\
&C_{act} : \{0,1\}^n \to \{0,1\}^|C| \\
&C_{act}(h) \equiv (h(c_0), h(c_1), h(c_2), \dots h(c_n)) \\
&\quad = (h^T c_0, h^T c_1, \dots, h^T c_n) \\
&\quad = h^T (c_0, c_1, \dots c_n)
\end{aligned}$$

If we regard $(c_0, c_1, c_2, \dots c_n) \subseteq \mathbb{R}^{n \times |C|}$ as a matrix, then we can see that $C_{act}$ is a *linear function.*

Now, if the set $C$ shatters $\mathcal{H}$, then:

1 The function $C_{act}$ will produce every element in $\{0,1\}^{|C|}$

2 the function $C_{act}$ will have full image.

3 This is only possible when the dimension of the domain is less than or equal to the dimension of the range.

4 the largest set that can be shattered is the largest matrix $C \subseteq R^{n \times |C|}$ such that the function $C_{act}$ has full range.

5 Thus, $|C| \leq n$ for $C_{act}$ to have full range.

6 We can achieve $|C| = n$ by picking $C = I_{n \times n}$. In other words, the element $c_i$ will be the $i$th row of the identity matrix. That is $c_i[j] = \mathbb{1}[i = j] = \begin{cases} 1 & \text{i = j} \\ 0 & \text{otherwise} \end{cases}$. Clearly, this $C$ is shattered since the function $C_{act}$ is the identity function which will produce every single output in $\{0,1\}^{|C|} = \{0,1\}^n = \{0,1\}^{|\mathcal{H}|}$.

7 $|C| = n$ is the largest possible, since $C_{act}$ is a function, and the size of its image is at most the size of the domain. Since the domain $\mathcal{H}$ has $2^n$ elements, the image too can have at most $2^n$ elements, which it does when $|C| = n$, since $|0, 1^{|C|}| = 2^{|C|}$.

7.5 $|C| = n$ is the largest possible, since $C_{act}$ is linear. For a linear function to be surjective, we need $Dim(domain) Dim(range)$. Hence, $Dim(Domain) = Dim(\mathcal{H}) = n \geq Dim(range) = Dim(\{0,1\}^{|C|}) = c$. That is, $n \geq |C|$.

Hence, we conclude that $\text{VCdim}(\mathcal{H}) = n$.

## 6.8, Q5:

Let $\mathcal{H}^d$ be the class of axis-aligned bounding boxes in $\mathbb{R}^d$. Show that the VC dimensions of $\mathcal{H}^d$ is $2d$.

### Solution

Formally, we have

$$h_{\vec{l},\vec{r}}(\vec{p}) \equiv \begin{cases} 1 & l[i] \leq p[i] \leq r[i] \text{ for all } i \in \{1, 2, \dots, n\} \\ 0 & \text{otherwise} \end{cases} \qquad \mathcal{H}^d \equiv \left\{ h_{\vec{l},\vec{r}} : \mathbb{R}^d \to \{0,1\} \mid \forall \vec{l}, \vec{r} \in \mathbb{R}^d \right\}$$

We claim that the set of points:

$$S \equiv S^{+} \cup S^{-}$$
$$S^{+} \equiv \{p[i] \equiv 1, p[j \neq i] \equiv 0 : i \in [d], p \in \mathbb{R}^{d}\}$$
$$S^{-} \equiv \{p[i] \equiv -1, p[j \neq i] \equiv 0 : i \in [d], p \in \mathbb{R}^{d}\}$$

shatters the hypothesis class $\mathcal{H}^{d}$.

We first show that $|H^{d}|$ restricted to $S$ expresses all functions $S \to \{0, 1\}$.

We will proceed by induction on the dimension $n$. In the case of ($n = 2$), we have already shown this as part of the course (shattering of 4 points by rectangles). We assume that this possible for $n = d - 1$. We need to show that this possible for $n = d$.

Let us say that we are currently trying to shatter a set $T$. If the points in $T$ lie in a $(d - 1)$ subspace of

To show that a set of size $2d + 1$ cannot be shatt