# CMU 10-715: Homework 4

VC Dimension

**DUE: Oct. 10, 2020, 11:59 PM.**

## Instructions:

- **Collaboration policy:** Collaboration on solving the homework is allowed, after you have thought about the problems on your own. It is also OK to get clarification (but not solutions) from books, again after you have thought about the problems on your own. Please don't search for answers on the web, previous years' homeworks, etc. (please ask the TAs if you are not sure if you can use a particular reference). There are two requirements: first, cite your collaborators fully and completely (e.g., "Alice explained to me what is asked in Question 4.3"). Second, write your solution *independently*: close the book and all of your notes, and send collaborators out of the room, so that the solution comes from you only.
- **Submitting your work:** Assignments should be submitted as PDFs using Gradescope unless explicitly stated otherwise. Each derivation/proof should be completed on a separate page. Submissions can be handwritten, but should be labeled and clearly legible. Else, submission can be written in LaTeX.
- **Late days:** For each homework you get three late days to be used only when anything urgent comes up. No points will be deducted for using these late days. We will consider an honor system where we will rely on you to use the late days appropriately.

# 1 VC Dimension of Polynomial Classifiers [100]

We discussed in class that the VC dimension of a single segment line classifier in $\mathbb{R}$ is 2. In this homework we are going to analyze the VC dimension of a similar hypthosesis class, parametrizing it with a well known family of functions.

Let $M = \{x_1, x_2, ..., x_m\}$ be a set of $m$ points in $\mathbb{R}$ such that $x_i \leq x_j$ if $i < j$ in $\mathbb{R}$ and the set $\{I_0, I_1, I_{m+1}\}$ of m+1 intervals associated with $M$, where:

$$I_0 = (-\infty, x_1], I_1 = (x_1, x_2], I_2 = (x_2, x_3], ..., I_m = (x_m, \infty).$$

Now consider the classifier $h_M$ which assigns the same label (+1 or -1) for all points inside any interval. Formally, if $x_1, x_2 \in I_j$ with $j \in \{0, ..., m\}$, then $h_M(x_1) = h_M(x_2)$ and $h_M(x_1) \in \{-1, 1\}$.

Let $\mathcal{H}_m = \{h_M | \quad |M| = m\}$, the hypothesis class consisting of all classifiers $h_M$ with $M$ consisting of $m$ points in $\mathbb{R}$.

Consider the set of polynomials functions of $x \in \mathbb{R}$ with degree at most $d$,

$$P_d(x) = \sum_{i=0}^{d} a_i x^i$$

with $a_i \in \mathbb{R}$. For a given $P_d(x)$ we define the classifier $f_d(x)$ as,

$$f_d(x) = \begin{cases} +1 & \text{if } P_d(x) \geq 0 \\ -1 & \text{otherwise.} \end{cases}$$

(a) (10 points) We are going to compute the VC dimension of the hypothesis class defined by $f_d$ below, but before doing so it will be useful to develop the intuition on the problem. How is the polynomial classifier $f_d$ related to they hypothesis class $\mathcal{H}_m$ defined above?

(b) (15 points) Find a function $\phi: \mathbb{R} \mapsto \mathbb{R}^{d+1}$ that allows you to write the polynomial $P_d(x) = \langle a, \phi(x) \rangle$ as a dot product between $a \in \mathbb{R}^{d+1}$ and $\phi(x) \in \mathbb{R}^{d+1}$.

(c) (30 points) Think of the implication of expressing the polynomials as a linear function in $\mathbb{R}^{d+1}$ to prove that the VC dimension of the decision function $f_d$ is at most $d+1$.

(d) (15 points) To prove that the VC dimension is grater than or equal to $d+1$, first show that there exists a set of points $\{x_1, x_2, \ldots, x_{d+1}\} \subseteq \mathbb{R}$ such that $\{\phi(x_1), \phi(x_2), \ldots, \phi(x_{d+1})\}$, are $d+1$ linearly independent vectors in $\mathbb{R}^{d+1}$.

(d) (15 points) Then show that with the vectors from part (d) of the problem and your findings from part (b) of the problem you can shatter any possible labeling vector $y = (y_1, y_2, \ldots, y_{d+1})^\intercal$ with $y_i \in \{-1, +1\}$. What is the implied $f_d$ classifier?

(e) (15 points) Using (c) and (d) infer the VC dimension of the $d-$th polynomial classifier $f_d$. What can you say about the VC dimension of the family of all possible polynomial classifiers on the real numbers?