2.4 Exercises

Karl Hiner

September 3, 2023

1 Overfitting of polynomial matching

We have shown that the predictor defined in Equation (2.3) leads to overfitting. While this predictor seems to be very unnatural, the goal of this exercise is to show that it can be described as a thresholded polynomial. That is, show that given a training set

$$S = \{(\boldsymbol{x}_i, f(\boldsymbol{x}_i))\}_{i=1}^m \subseteq (\mathbb{R}^d \times \{0, 1\})^m,$$

there exists a polynomial p_S such that $h_S(\boldsymbol{x}) = 1$ iff $p_S(\boldsymbol{x}) \geq 0$, where h_S is as defined in Equation (2.3):

$$h_S(\boldsymbol{x}) = \begin{cases} y_i & \text{if } \exists i \in [m] \text{ s.t. } \boldsymbol{x}_i = \boldsymbol{x}, \\ 0 & \text{otherwise} \end{cases}.$$

It follows that learning the class of all thresholded polynomials using the ERM rule may lead to overfitting.

Answer: We are tasked with showing that, given any training set S mapping a set of real vectors to corresponding boolean scalars, we can construct a polynomial p_S such that $h_S(\boldsymbol{x}) = 1$ iff $p_S(\boldsymbol{x}) \geq 0$.

Here, h_S is a "bad" hypothesis that (maximally) overfits the training set S by simply acting as a lookup table, returning 1 when it receives a value $\boldsymbol{x}_i \in S$ with $f(\boldsymbol{x}_i) = y_i = 1$, and returning 0 otherwise.

In order to construct a polynomial with the same behavior, we must ensure the polynomial crosses the x-axis only at the points where $y = 1$.

First, we consider one term of the polynomial for a single point $(\boldsymbol{x}_i, y_i) \in S$, where $y_i = f(\boldsymbol{x}_i)$. We want the term to return zero if $\boldsymbol{x} \neq \boldsymbol{x}_i$, and y_i if $\boldsymbol{x} = \boldsymbol{x}_i$. Here is such a term, $t_i(\boldsymbol{x})$:

$$t_i(\boldsymbol{x}) = y_i \prod_{j=1}^d (x_j - (\boldsymbol{x}_i)_j)^2.$$

By taking the square of the difference between the j_{th} component of $\boldsymbol{x}$ and $\boldsymbol{x}_i$ for each j and multyplying these together, this term will be zero if $\boldsymbol{x} \neq \boldsymbol{x}_i$ in any dimension.

Then, we can express the full polynomial $p_S(\boldsymbol{x})$ as the sum of each term:

$$p_S(\boldsymbol{x}) = \sum_{i=1}^m t_i(\boldsymbol{x}) = \sum_{i=1}^m y_i \left(\prod_{j=1}^d (x_j - (\boldsymbol{x}_i)_j)^2 \right).$$

This polynomial will be positive at each point $\boldsymbol{x}_i$ where $y_i = 1$ and zero where $y_i = 0$. It will also be zero for points not in S, and so we have $h_S(\boldsymbol{x}) = 1$ iff $p_S(\boldsymbol{x}) \geq 0$.

TODO show with example.

2

Let $\mathcal{H}$ be a class of binary classifiers over a domain $\mathcal{X}$. Let $\mathcal{D}$ be an unknown distribution over $\mathcal{X}$, and let f be the target hypothesis in $\mathcal{H}$. Fix some $h \in \mathcal{H}$. Show that the expected value of $L_S(h)$ over the choice of $S|_x$ equals $L_{(\mathcal{D},f)}(h)$, namely,

$$\mathop{\mathbb{E}}_{S|_x \sim \mathcal{D}^m} [L_S(h)] = L_{(\mathcal{D},f)}(h).$$

Answer:

Want to show:

$$\mathop{\mathbb{E}}_{S|_x \sim \mathcal{D}^m} [L_S(h)] = L_{(\mathcal{D},f)}(h).$$

3 Axis aligned rectangles

An axis aligned rectangle classifier in the plane is a classifier that assigns the value 1 to a point iff it is inside a certain rectangle. Formally, given real numbers $a_1 \leq b_1, a_2 \leq b_2$, define the classifier $h_{(a_1,b_1,a_2,b_2)}$ by

$$h_{(a_1,b_1,a_2,b_2)}(x_1, x_2) = \begin{cases} 1 & \text{if } a_1 \leq x_1 \leq b_1 \text{ and } a_2 \leq x_2 \leq b_2, \\ 0 & \text{otherwise} \end{cases}.$$

The class of all axis aligned rectangles in the plane is defined as

$$\mathcal{H}^2_{\text{rec}} = \{h_{(a_1,b_1,a_2,b_2)} : a_1 \leq b_1, a_2 \leq b_2\}.$$

Note that this is an infinite size hypothesis class. Throughout this exercise we rely on the realizability assumption.