Assigned reading: Chapters 3-6 of Shalev-Shwartz and Ben-David, Understanding Machine Learning, from Theory to Algorithms.

On the VC dimension of thresholded sums of functions.

Let $\text{pos}$ map a vector $v$ to the binary vector of the same dimension, such that $\text{pos}(v)_i = 1_{\{v_i \geq 0\}}$. Similarly, given a function $g : \mathbb{Z} \rightarrow \mathbb{R}$ let $\text{pos}(g) : \mathbb{Z} \rightarrow \{0, 1\}$ be defined by $\text{pos}(g)(z) = \text{pos}(g(z))$, and if $\mathcal{G}$ is a family of such functions let $\text{pos}(\mathcal{G}) = \{\text{pos}(g) : g \in \mathcal{G}\}$. Throughout the remainder of this problem, suppose that $\mathcal{G}$ is a linear space of functions on $\mathbb{Z}$.

1. Let $m$ be a finite positive integer. Show that the rank of $\mathcal{G}$ is $m$ if and only if there are $m$ linearly independent functions $\psi_1, \ldots, \psi_m$ in $\mathcal{G}$ such that any $g \in \mathcal{G}$ can be represented as $g(z) = \sum_{i=1}^{m} c_i \psi_i(z)$ for all $z \in \mathbb{Z}$. (Hint: The set of spanning functions can be selected greedily.)

2. Suppose the rank of $\mathcal{G}$ is finite and equal to $m$, and suppose $\psi_1, \ldots, \psi_m$ are as in part (a). Show that there exists $\{z_1, \ldots, z_m\}$ such that the $m \times m$ matrix $(\psi_i(z_j))_{1 \leq i,j \leq m}$ has full rank. (You can skip proving this part and use it for the next part.)

3. Suppose the rank of $\mathcal{G}$ is finite and equal to $m$. Show that $V(\text{pos}(\mathcal{G})) \geq m$. Here, $V$ denotes VC dimension applied to sets of binary valued functions, which are equivalent to sets of subsets. (Hint: Use part 2.)

4. (VC dimension of Dudley classes) Suppose the rank of $\mathcal{G}$ is finite and equal to $m$, and let $h : \mathbb{Z} \rightarrow \mathbb{R}$, with $h$ not necessarily in $\mathcal{G}$. Show that $V(\text{pos}(\mathcal{G} + h)) \geq m$.

5. Consider the case where $\mathcal{C}$ is the class of axis-parallel (hyper) cubes in $\mathbb{R}^d$. Show that $V(\mathcal{C}) \leq d + 1$. (Hint: Use the previous part, and you can assume the stronger result that $V(\text{pos}(\mathcal{G} + h)) = m$.)