2. Expected Risk Minimization and Abstract Tools for Uniform Approximation

**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 $g: \mathbb{Z} \to \mathbb{R}$ let $\text{pos}(g): \mathbb{Z} \to \{0, 1\}$ 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 $\mathbb{Z}$.

A set of real-valued functions $\psi_1, \ldots, \psi_m$ on $\mathbb{Z}$ is said to be linearly independent if the only vector $(c_i : i \in [k])$ such that $\sum_{i=1}^m c_i \psi_i \equiv 0$ on $\mathbb{Z}$ is the zero vector. The (linear) rank of $\mathcal{G}$ is the supremum of all $k$ such that there exist $k$ linearly independent elements of $\mathcal{G}$.

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. (Hint: Think greedy again!)

3. Suppose the rank of $\mathcal{G}$ is finite and equal to $m$. Show that $V(\text{pos}(\mathcal{G})) = m$. Here, $V$ denotes VC dimension applied to sets of binary valued functions, which are equivalent to sets of subsets. (Hint: Since it is shown in the notes that $V(\text{pos}(\mathcal{G})) \leq m$, you only need to establish the reverse inequality. Use part (b).)

4. (VC dimension of Dudley classes) Suppose the rank of $\mathcal{G}$ is finite and equal to $m$, and let $h : \mathbb{Z} \to \mathbb{R}$, with $h$ not necessarily in $\mathcal{G}$. Show that $V(\text{pos}(\mathcal{G} + h)) = m$. (Hint: To show $V(\text{pos}(\mathcal{G} + h)) \leq m$, explain how the proof in the notes can be modified to handle nonzero $h$. For the reverse direction, a modification of part (c) suffices.)

5. Consider the example in the notes, such that $\mathcal{C}$ is the class of closed balls in $\mathbb{R}^d$. Show that $V(\mathcal{C}) \leq d + 1$. (Hint: Use previous part.)