# AlgoEcon Final Project

May 2023

## 1 Natarajan Dimension

In problem 5 we are given the hypothesis class, $\mathcal{H}' = \{h'_\tau : \tau \in \mathbb{R}\}$ , where $h'_\tau$ applies the threshold rule element-wise to the input, i.e. for $\vec{x} \in \mathbb{R}$ and $\vec{y} = h'_\tau(\vec{x})$; then $\vec{y}(i) = 1 \Leftrightarrow \vec{x}(i) \geq \tau, i \in \{1, ...n\}$.

Our definition of shattering in the multi class setting is given as follows:

A set $S$ of vectors in $\mathbb{R}^n$ is shattered by the hypothesis class $\mathcal{H}'$ if the follow conditions are satisfied:

1. $\forall \vec{x}_k \in S$, $\exists \vec{x}_{k,A}, \vec{x}_{k,B} \in \{0,1\}^n$, such that $\vec{x}_{k,A} \neq \vec{x}_{k,B}$.
2. $\forall \phi \in \Phi$ where $\Phi = \{\phi \in \mathcal{F} : \phi : \{\vec{x}_k \in S\} \to \{\vec{x}_{k,A}, \vec{x}_{k,B}\}\}$ $\exists h'_\tau \in \mathcal{H}'$ such that $\forall \vec{x}_k \in S$ $h'_\tau(\vec{x}_k) = \phi(\vec{x}_k)$, (I've denoted $\Phi$ this way to suggest that each $\vec{x}_k$ need not be mapped to the same co-domain).

The Natarajan dimension is then given by the cardinality of the largest set $\mathcal{H}'$ can shatter. The proof proceeds by first showing that $\mathcal{H}'$ cannot shatter any $|S| = 2$ and second, arguing that a hypothesis class cannot shatter any set larger than the smallest set it cannot shatter.

1. **Claim**: $\mathcal{H}'$ cannot shatter any set $S$ of two vectors $\vec{x}_1, \vec{x}_2 \in \mathbb{R}^2$, $|S| = 2$.

   **Proof:** We proceed by contradiction, showing that no $h'_\tau$ can reproduce all $\phi$ for an arbitrary choice of vectors $S = \{\vec{x}_1, \vec{x}_2\}$. Suppose condition 1 holds. Take $\{\vec{x}_{1,1}, \vec{x}_{1,2}\}$, $\vec{x}_{1,1} \neq \vec{x}_{1,2}$, and $\{\vec{x}_{2,1}, \vec{x}_{2,2}\}$, $\vec{x}_{2,1} \neq \vec{x}_{2,2}$ to be the sets of arbitrary vectors in $\{0,1\}^n$ associated with $\vec{x_1}$ and $\vec{x_2}$ respectively. Since, $\vec{x}_{1,1} \neq \vec{x}_{1,2}$ and $\vec{x}_{2,1} \neq \vec{x}_{2,2}$, $\exists i, j \in \{1, 2\}$ such that $\vec{x}_{1,1}(i) \neq \vec{x}_{1,2}(i)$ and $\vec{x}_{2,1}(j) \neq \vec{x}_{2,2}(j)$ where here $\vec{x}_{1,1}(i)$ denotes the $i^{th}$ index of $\vec{x}_{1,1}$, etc. WLOG take $\vec{x}_{1,1}(i) = \vec{x}_{2,1}(j) = 0$, $\vec{x}_{1,2}(i) = \vec{x}_{2,2}(j) = 1$, and $\vec{x_1}(i) \leq \vec{x_2}(j)$ $(*)$. Now suppose condition 2 holds. This implies that $\forall \phi \in \Phi_2$, $\exists h'_\tau \in \mathcal{H}'$ such that both $\phi(\vec{x_1}) = h'_\tau(\vec{x_1})$ and $\phi(\vec{x_2}) = h'_\tau(\vec{x_2})$ where in this regime, $\Phi_2 = \{\phi \in \mathcal{F} : \phi : \{\vec{x}_k \in S\} \to \{\vec{x}_{k,1}, \vec{x}_{k,2}\}\}$. Now take a particular $\phi_0 \in \Phi_2$ such that $\phi_0(\vec{x}_1) = \vec{x}_{1,2}$ and $\phi_0(\vec{x}_2) = \vec{x}_{2,1}$. Since we are assuming condition 2 holds we must have $h'_\tau$ such that $h'_\tau(\vec{x}_1) = \vec{x}_{1,2}$ and $h'_\tau(\vec{x}_2) = \vec{x}_{2,1}$. However, since $\vec{x}_{2,1}(j) = 0$ this implies $\vec{x}_2(j) < \tau$. Transitively, we achieve $\vec{x}_1(i) < \vec{x}_2(j) \Rightarrow \vec{x}_1(i) < \tau$ and hence by definition $h'_\tau(\vec{x}_1)(i) = \vec{x}_{1,2}(i) = 0$. Comparing with the first generalization in $(*)$, we arrive at a contradiction. Since our choices of $\vec{x}_1, \vec{x}_2$, the corresponding $\vec{x}_{1,1}, \vec{x}_{1,2}, \vec{x}_{2,1}, \vec{x}_{2,2}$, and $i, j$ are general up to relabelling, we have shown that if condition 1 holds then condition 2 must necessarily fail for an arbitrary choice of $S$. Thus, $\mathcal{H}'$ cannot shatter any $S$, $|S| = 2$.

2. **Claim**: Any hypothesis class $\mathcal{H}$ which cannot shatter any set $S$, $|S| = m$, cannot shatter any set $T$, $|T| = l$, $l > m$, where $S$ and $T$ are subsets of some feature space $\mathcal{X}$.

   **Proof:** This is an informal proof of the contra-positive. Suppose $\mathcal{H}$ shatters $T$. Then take $S = T \backslash U$ where $U \subset T, U \neq \emptyset$. Since $\mathcal{H}$ shatters $T$ by definition $\forall \vec{x}_k \in T$, $\exists \vec{x}_{k,A}, \vec{x}_{k,B} \in \{0,1\}^n$, such that $\vec{x}_{k,A} \neq \vec{x}_{k,B}$ and $\forall \phi \in \Phi$ where $\Phi = \{\phi \in \mathcal{F} : \phi : \{\vec{x}_k \in T\} \to \{\vec{x}_{k,A}, \vec{x}_{k,B}\}\}$ $\exists h'_\tau \in \mathcal{H}'$ such that $\forall \vec{x}_k \in T$ $h'_\tau(\vec{x}_k) = \phi(\vec{x}_k)$. Take an arbitrary $\vec{x} \in S$. Since $S \subset T \Rightarrow \vec{x} \in T$, but $\vec{x} \in T$ implies $\vec{x}$ satisfies the previous two shattering conditions. Since our choice of $\vec{x}$ is arbitrary, it is clear that all $\vec{x} \in S$ satisfy the shattering conditions. Hence, $\mathcal{H}$ shatters $S$.

Combining the two proofs we have that $\mathcal{H}'$ cannot shatter any set $S$ of vectors in $\mathbb{R}^n$, $|S| \geq 2$. Showing that $H'$ can shatter a set $S$, $|S| = 1$, reduces to problem 1. Hence, the $Ndim(\mathcal{H}') = 1$.

# 2 Bonus

**Claim:** The $Ndim(\hat{\mathcal{H}}) = n$

**Proof:** To prove, we need only to show that $\hat{\mathcal{H}}$ shatters a single set $S$, $|S| = n$, of vectors in $\mathbb{R}^n$. To do this, we will show that, $\mathcal{H}^\dagger \subset \hat{\mathcal{H}}$, $\mathcal{H}^\dagger = \{\hat{h_\Gamma} \in \hat{\mathcal{H}} : \forall i \in \{1, .., n\}, \Gamma_i > 0\}$, shatters such a set. For ease of calculation we will take $S$ to be the standard set of unit vectors which form an orthonormal basis for $\mathcal{R}^n$. We proceed by induction on $m$ up to $n$:

**Base Case:** Take $n = 1$, $S = \{\vec{e_1} = 1\}$, $1 \in \mathbb{R}$. We have only two vectors with which 1 can be associated, namely $0, 1$. Likewise, there are only two mappings $\phi$, $1 \to 0$, $1 \to 1$. For each mapping we can generate a corresponding $\hat{h}_\Gamma$, simply by setting $\Gamma_1 = \Gamma < 1$ and $\Gamma \geq 1$ respectively. We are done.

**Induction:** Suppose $\mathcal{H}^\dagger$ shatters some set $T_m = \{\vec{e_1}, \ldots, \vec{e_m}\}$ and $T_m \subseteq S$, $|T_m| = m \leq n$, where $\forall \vec{e_k} \in T_m$, $\vec{0}$, $\vec{e_k}$ are the vectors associated with $\vec{e_k}$. By our inductive hypothesis, we assume $\forall \phi \in \Phi_m = \{\phi \in \mathcal{F} : \phi : \{\vec{e_k} \in T_m\} \to \{\vec{0}, \vec{e_k}\}\}$ $\exists h^\dagger_\Gamma \in \mathcal{H}^\dagger$ such that $\forall \vec{e_k} \in T_m$, $h^\dagger_\Gamma(\vec{e_k}) = \phi(\vec{e_k})$. Notice, if $T_m = S$, then we are done. Else, $T_m \neq S \Rightarrow T_m \subset S$ and $m < n$. $T_m \subset S \Rightarrow \exists \vec{e_j} \in S$ such that $\vec{e_j} \notin T$. We define a new set $T_{m+1} = T_m \cup \{\vec{e_j}\}$ where $|T_{m+1}| = m + 1 \leq n$. Clearly, $T \subseteq S$. Let us introduce a new set of functions $\Phi_{m+1} = \{\phi \in \mathcal{F} : \phi : \{x_k \in T_{m+1}\} \to \{\vec{0}, \vec{e_k}\}\}$. Take $\phi_o \in \Phi_{m+1}$. Now define $\phi_r = \phi_o|_{T_m}$, where $\phi_r$ is the restriction of $\phi_o$ to the domain of $T_m$. Note, $\phi_r = \phi_o|_{T_m} \Rightarrow \forall \vec{e_k} \in S$, $\phi_o(\vec{e_k}) = \phi_r(\vec{e_k}) \Rightarrow \phi_r \in \Phi_m$ (since otherwise we'd have a contradiction). Further, $\phi_r \in \Phi_m \Rightarrow \exists h^\dagger_{\Gamma^*} \in \mathcal{H}^\dagger$ such that $\forall \vec{e_k} \in T_m$ $h^\dagger_{\Gamma^*}(\vec{e_k}) = \phi_r(\vec{e_k}) = \phi_o(\vec{e_k})$. Since our choice of $\phi_o$ was arbitrary, we have that $\forall \phi_o \in \Phi_{m+1} \exists h^\dagger_{\Gamma^*}$ such that $\forall \vec{e_k} \in T_m$, $h^\dagger_{\Gamma^*}(\vec{e_k}) = \phi_o(\vec{e_k})$. If we can show that, for a particular choice of $\Gamma_j \in \Gamma^*$, $h^\dagger_{\Gamma^*}(\vec{e_j}) = \phi_o(\vec{e_j})$ and still retain $\forall \vec{e_k} \in T_m$ $h^\dagger_{\Gamma^*}(\vec{e_k}) = \phi_o(\vec{e_k})$ then we will have proven $\mathcal{H}^\dagger$, and therefore $\hat{\mathcal{H}}$, shatters $T_{m+1}$. First we show that the action of $h^\dagger_{\Gamma^*}$ on $T_m$ is invariant for a change of $\Gamma^* \to \Gamma'$ by varying only $\Gamma^*_j$. That is $\forall \Gamma^*, \Gamma' \in \mathbb{R}^n$, $\forall i \in \{1, .., n\}, i \neq j$ $\Gamma^*_i = \Gamma'_i$, $\forall \vec{e_k} \in T_m$, $h^\dagger_{\Gamma^*}(\vec{e_k}) = h^\dagger_{\Gamma'}(\vec{e_k})$. Suppose not: this implies $\exists \vec{e_k} \in T_m$ such that $h^\dagger_{\Gamma^*}(\vec{e_k}) \neq h^\dagger_{\Gamma'}(\vec{e_k})$. We know $\forall i \neq j$, $\Gamma^*_i = \Gamma'_i \Rightarrow h^\dagger_{\Gamma^*}(\vec{e_k})_i = h^\dagger_{\Gamma'}(\vec{e_k})_i$. Now note that $h^\dagger_{\Gamma^*}(\vec{e_k})_j = \phi_r(\vec{e_k})_j = 0$; this is because $\phi_r(\vec{e_k})$, and therefore $h^\dagger_{\Gamma^*}(\vec{e_k})$, are in $\{\vec{0}, \vec{e_k}\}$, but $k \neq j$ since $\vec{e_j}$ is defined such that $\vec{e_j} \notin T_m$ and all of $\vec{e_l} \in S$ distinct from the definition of a basis. Likewise, $h^\dagger_{\Gamma'}(\vec{e_k})_j = 0$ since $\vec{e_k}_j = \delta_{k,j} = 0$, (where here we've used the definition of the standard set of unit vectors that form a basis for $\mathcal{R}^n$), and $h^\dagger_{\Gamma'} \in \mathcal{H}^\dagger \Rightarrow \Gamma' > 0 \Rightarrow (\vec{e_k})_j \ngeq \Gamma'$. Thus we have achieved $\nexists i$ such that $h^\dagger_{\Gamma^*}(\vec{e_k})_i \neq h^\dagger_{\Gamma'}(\vec{e_k})_i$, a contradiction. Now we are ready to complete the proof. Consider our arbitrary choice $\phi_o \in \Phi_{m+1}$. Take $h^\dagger_{\Gamma^*}$ as before. Suppose $\phi_o(\vec{e_j}) = \vec{0}$, then we take $h^\dagger_{\Gamma'}$, where $i \neq j, \Gamma'_i = \Gamma^*_i, \Gamma'_j = 1$. If instead, $\phi_o(\vec{e_j}) = \vec{e_j}$, we take $\Gamma'_j = \frac{1}{2}$. In both cases, we get that $h^\dagger_{\Gamma'}(\vec{e_j}) = \phi_o(\vec{e_j})$ and by the invariance lemma above we retain $h^\dagger_{\Gamma'}(\vec{e_k}) = \phi_o(\vec{e_k}), \vec{e_k} \in T_m$. Thus, $\forall phi \in \Phi_{m+1}$ $\exists h^\dagger_{\Gamma'} \in \mathcal{H}^\dagger$ such that $\forall \vec{e_k} \in T_{m+1}$ $h^\dagger_{\Gamma'}(\vec{e_k}) = phi(\vec{e_k})$ or $\mathcal{H}^\dagger$ and therefore $\hat{\mathcal{H}}$ shatters $T_{m+1}$.

Our assumption that $|T| = m \leq n$ holds up to $m = n$. So by the principle of induction up to $n$ we have shown for any $m \leq n$ there exists a set $S$, $|S| = m$ of unit vectors such that $\mathcal{H}^\dagger$ and therefore $\hat{\mathcal{H}}$ shatters $S$. Taking $m = n$ we see that this implies $Ndim(\hat{\mathcal{H}}) = n$