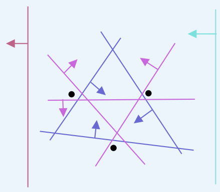
Problem 1. Define $\mathcal{R}$ to be the set of all halfplanes in $\mathbb{R}^2$. Prove: The VC-dimension of the range space $(\mathbb{R}^2, \mathcal{R})$ is 3.

Proof.

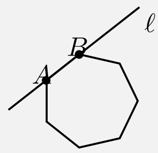
(a) The following diagram illustrates a scattered set $\{(0,-1),(-2,0),(0,1)\}$ with size 3.



(b) By Radon's Theorem, there exists a partition of $\mathcal{Y} \cup \mathcal{Z} = P$ and $\mathcal{Y} \cap \mathcal{Z} = \emptyset$ such that their convex hull intersect. Without loss of generality, we assume that $\mathcal{Y} = \{p_1, p_2\}, \mathcal{Z} = \{p_3, p_4\}$. Clearly, the convex hulls of $\mathcal{Y}, \mathcal{Z}$ are both line segments. Now we claim that $\mathcal{Y}$ cannot be shattered by any $h \in \mathcal{R}$. Suppose the otherwise that such h exists, then all points from $\mathcal{Y}$ must lie on one side of the halfplane, and points from $\mathcal{Z}$ on the other side. Now, $\mathcal{CH}(\mathcal{Y}) \subseteq h$, $\mathcal{CH}(\mathcal{Z}) \subseteq \overline{h}$, by convexity. This contradicts the fact that $\mathcal{CH}(\mathcal{Y}) \cap \mathcal{CH}(\mathcal{Z}) \neq \emptyset$.

We therefore conclude that for the range space $\mathcal{S} = (\mathbb{R}^2, \mathcal{R})$, $\dim_{\mathrm{VC}}(\mathcal{S}) = 3$. □

Problem 2. Let $\mathcal{R}$ be the set of all triangles in $\mathbb{R}^2$. Let S be an arbitrary set of lines in $\mathbb{R}^2$. Given any triangle $q \in \mathcal{R}$, define $q(S)$ to be the set of lines that intersect q. In addition define $\mathcal{R}_S = \{q(S) | q \in \mathcal{R}\}$. Let G be a regular 7-gon. Let S be the set of supporting lines of the 7 edges of G (the supporting line ℓ of the edge $\overline{AB}$ is shown below). Prove: $\mathcal{R}_S = 2^S$.



Proof. Let $S = \{\ell_1, \ell_2, \ldots, \ell_7\}$, $T = \{t_1, t_2, \ldots, t_7\}$ be the intersection of $\ell_1 \cap \ell_2, \ell_2 \cap \ell_3, \ldots, \ell_7 \cap \ell_2$, and $M = \{m_1, m_2, \ldots, m_7\}$ be the midpoints of the segments of the 7-gon having $\ell_1, \ell_2, \ldots, \ell_7$ as supporting lines, respectively. Consider a triangle q with vertices from T, M, or interior of G. This triangle is constrainted by the 7-gon. Therefore, $q(S) = \{\ell_i, \ell_{i\%7+1} | m_i \in q\} \cup \{\ell_i | v_i \in q\}$.

Now, we can scatter any $S' \subseteq S$ with $Q \in \mathcal{R}$ in the following manner:

If $|S'| = 7$, then a triangle covering the 7-gon suffices to intersect all lines. If $|S'| = 6$, then all 3 sides are picked from T each corresponding to 2 lines in S'. If $|S'| = 4$, there should be at least one i with $\ell_i, \ell_{i\%7+1} \in S'$. As such, we pick two elements from M and one element from T. If $|S'| \leq 3$, then we select $3 - |S'|$ non-overlapping points from interior of G, and $|S'|$ points from M each corresponding to a line in S'. For $|S'| = 5$, if the lines selected are continuous piece of segments around the polygon, then this is easy. The other case to consider is having the polygon broken down into two pieces and have pieces with size 1 and 4 or 2 and 3. This is also easy. □

Problem 3. Let $(X, \mathcal{F})$ be a set system with VC-dimension λ. Fix an arbitrary element $e \in X$. Define

$$\begin{aligned} X' &= X \backslash \{e\} \\ \mathcal{F}'' &= \{Q \backslash \{e\} | Q \in \mathcal{F}, e \in Q \text{ and } Q \backslash \{e\} \notin \mathcal{F}\} \\ \mathcal{F}''' &= \{Q | Q \in \mathcal{F} \text{ and } e \notin Q\} \\ \mathcal{F}_2 &= \mathcal{F}'' \cup \mathcal{F}''' \end{aligned}$$

Prove: the VC-dimension of the set system $(X', \mathcal{F}_2)$ is at most λ.

Proof. Suppose the otherwise that the statement is false. Then,

$$\exists S' \subseteq X', \text{ s.t. } |S'| = \lambda \text{ and } S' \text{ can be shattered by } \mathcal{F}_2.$$

Claim: S' can be shattered by $\mathcal{F}$, with $|S'| = \lambda + 1$, this leads to a contradiction as $S' \subseteq X$ has VC dimension $\lambda + 1$.

Our goal is to prove that

$$\forall T \subseteq S', \exists Q \in \mathcal{F} \text{ s.t. } T = Q \cap S'.$$

As S' can be shattered by $\mathcal{F}_2$, it implies that

$$\exists Q' \in \mathcal{F}_2 \text{ s.t. } T = Q' \cap S'.$$

It suffices to consider two cases: $Q' \in \mathcal{F}''$ and $Q' \in \mathcal{F}'''$.

(a) $Q' \in \mathcal{F}''' \Rightarrow Q' \in \mathcal{F}$, we can let $Q = Q'$.

(b) $Q' \in \mathcal{F}'' \Rightarrow (Q' \cup \{e\}) \cap S' = T$ (as $e \notin S'$). Now, by definition, it is clear that $Q' \cup \{e\} \in \mathcal{F}$, and we can let $Q = Q' \cup \{e\}$.

As we have found Q for every $T \subseteq S' \subseteq X$, the proof is completed. □

Problem 4. Fix $\varepsilon \in (0, 1)$. Let $(X, \mathcal{F})$ be any set system where X is a finite set of size n. Let Q be an arbitrary set in $\mathcal{F}$ with size at least εn. Suppose that we create a sample set $S \subseteq X$ by repeating the following operation $s \geq 1$ times on an initially empty S: take an element e from X uniformly at random and add e to S. Note that the final size of S may be less than s because it is possible for the same element to be taken more than once. Prove: $\mathbf{Pr}[S \cap Q = \emptyset] \leq 1/e^{\varepsilon \cdot s}$. Hint: Use the inequality $1 + x \leq e^x$.

Proof. $\mathbf{Pr}[S \cap Q \neq \emptyset] = \left(\frac{n - n \cdot \varepsilon}{n}\right)^s = (1 - \varepsilon)^s = (1 + (-\varepsilon))^s < 1 - \varepsilon s \leq e^{-\varepsilon \cdot s} = 1/e^{\varepsilon \cdot s}$. □

Problem 5. Fix $\varepsilon \in (0, 1)$. Let $(X, \mathcal{F})$ be any set system where X is a finite set of size n. Denote by λ the VC-dimension of $(X, \mathcal{F})$. Given a parameter $s \geq 1$, take a sample set S of X in the way described in Problem 4. Let B be the bad event where at least one $Q \in \mathcal{F}$ satisfies both conditions below:

- $|Q| \geq \varepsilon n$;
- $Q \cap S = \emptyset$.

Prove: $\mathbf{Pr}[B] \leq \Phi_\lambda(n)/e^{\varepsilon \cdot s}$.
Hint: Use the union bound in probability theory.
Theorem (Union Bound): Let $A_1, A_2, \ldots, A_k$ be arbitrary events where $k \geq 2$ (these events can be correlated to each other). Let $\cup_{i=1}^{k} A_i$ denote the event that at least one of the $A_1, A_2, \ldots, A_k$ occurs. It holds that

$$\mathbf{Pr}[\cup_{i=1}^{k} A_i] \leq \sum_{i=1}^{k} \mathbf{Pr}[A_i].$$

Proof. By Sauer's lemma, $|\mathcal{F}| \geq \Phi_\lambda(n)$. $\dim_{\mathrm{VC}}((X, \mathcal{F})) = \lambda$.

Define A_i for every $Q_i \in \mathcal{F}$, and set $A_i = 1$ if and only if $|Q| \geq \varepsilon \cdot n$ and $Q \cap S = \emptyset$.

$$\mathbf{Pr}\left[\bigcup_{i=1}^{k} A_i\right] \leq \sum_{i=1}^{\Phi_\lambda(n)} \mathbf{Pr}[A_i] \leq \Phi_\lambda(n)/e^{\varepsilon \cdot s}.$$

Note that in Problem 4 we derived that $\mathbf{Pr}[S \cap Q = \emptyset] = \mathbf{Pr}[A_i] \leq 1/e^{\varepsilon \cdot s}$. $\square$

Problem 6. Fix $\varepsilon \in (0,1)$. Let $(X, \mathcal{F})$ be any set system where X is a finite set of size n. Denote by λ the VC-dimension of $(X, \mathcal{F})$. Set $s = \lceil c \cdot \frac{\lambda}{\varepsilon} \log_2 n \rceil$ where c is a constant to be determined. Take a sample set S of X in the way described in Problem 4. Prove: when c is larger than a certain **constant**, the probability of S to be an ε-net of $(X, \mathcal{F})$ is at least $1 - 1/n^2$.
Hint: Use the fact that $\Phi_\lambda(n) = \sum_{i=0}^{\lambda} \binom{n}{i} < 1 + \lambda \cdot n^\lambda$ for any $\lambda \in [0, n]$. This inequality is very loose. The next statement gives a tighter result: for $\lambda \geq 2$, it holds that $\Phi_\lambda(n) \leq n^\lambda$.

Proof. ε**-net**: Given $S = (P, \mathcal{R}), \varepsilon > 0$, a subset $S \subseteq P$ is an ε-net if for any $Q \in \mathcal{R}$ with $\mu(Q) \geq \varepsilon$, then Q contains at least one point from S.

Essentially, we need to prove that $\Phi_\lambda(n)/e^{\varepsilon \cdot s} \leq 1/n^2$.

When $\lambda \geq 2$,

$$\begin{aligned}
\frac{\Phi_\lambda(n)}{e^{\varepsilon \cdot s}} &= \frac{n^\lambda}{e^{\varepsilon \cdot \lceil c \frac{\lambda}{\varepsilon} \log_2 n \rceil}} \\
&\leq \frac{n^\lambda}{e^{\varepsilon \cdot c \frac{\lambda}{\varepsilon} \log_2 n}} \\
&\leq \frac{n^\lambda}{e^{c\lambda \log_2 e \cdot \ln n}} \\
&\leq \frac{n^\lambda}{n^{c\lambda \log_2 e}} \\
&\leq n^{2(1 - c \log_2 e)}
\end{aligned}$$

Taking $c = 2$ satisfies the requirement $\frac{\Phi_\lambda(n)}{e^{\varepsilon \cdot s}} < n^{-3.7} < 1/n^2$.

$\lambda = 0$ implies $|Q| = 0$, as such, $\mathbf{Pr}[Q \cap S = \emptyset] = 1$.

$\lambda = 1$ implies $n \geq 1$. When $n = 1$, a sample with size $n = 1$ is the set X and thus it must be an ε-net. Otherwise, $\frac{\Phi_\lambda(n)}{e^{\varepsilon \cdot s}} \leq \frac{1+n}{e^{\varepsilon \cdot c \frac{\lambda}{\varepsilon} \cdot \log_2 n}} \leq \frac{2n}{n^{c \log_2 e}} \leq \frac{n^2}{n^{c \log_2 e}}$. Taking $c = 3$ satisfies the requirement.

$\square$

Problem 7. Fix $\varepsilon \in (0,1)$. Let $(X, \mathcal{F})$ be any set system where X is a finite set of size n. Denote by λ the VC-dimension of $(X, \mathcal{F})$. Prove: there is an ε-net of $(X, \mathcal{F})$ is with size $O(\frac{\lambda}{\varepsilon} \log n)$.
Hint: Use the result of Problem 6. Note that the sampling process in Problem 6 may fail to yield an ε-net with a positive probability.

Proof. $\lambda = 0$. No $Q \in \mathcal{F}$ can contain any element of X, otherwise we immediately have $\lambda = 1$. There ε-net with size 0 can work as required.

$\lambda = 1$. If $n = 1$, we have to include the only element $x \in X$, incurring an ε-net with size $O(1)$. Otherwise, by Problem 6, the probability that S with size $O(\frac{\lambda}{\varepsilon} \log n)$ is not an ε-net has probability smaller than $n^{-2} = 1$. ε with size $s = O(\frac{\lambda}{\varepsilon} \log n)$ has to exist.

$\lambda \geq 2$. If s is chosen such that $\mathbf{Pr}[B] < 1$, then the probability that S is an ε-net of X is positive. This implies that the ε with size $s = O(\frac{\lambda}{\varepsilon} \log n)$ has to exist. Now

$$\mathbf{Pr}[B] < \Phi_\lambda(n)/e^{\varepsilon \cdot s} \leq n^\lambda e^{-\varepsilon s} \leq 1 \Leftrightarrow n^\lambda \leq e^{\varepsilon s} \Leftrightarrow s \geq \frac{\lambda \ln n}{\varepsilon}.$$

Therefore, $s = O(\frac{\lambda}{\varepsilon} \log n)$. $\square$