Assignment 3 (ver 1)
(CS5061 Topics in Computing)

Deadline: 30th Nov.

1. Consider the problem of learning using halfspaces. Let the domain X be the Euclidean ball with radius r. i.e., $X = \{x \in \mathbb{R}^d \mid \|x\| \leq r\}$. Let the label set be $Y = \{-1, +1\}$ and the loss function $\ell(\vec{w}, (\vec{x}, y)) = \max\{0, 1 - y\langle\vec{w}, \vec{x}\rangle\}$ be the *hinge loss* function. Show that ℓ is R-Lipschitz.

2. Consider the setting of linear regression. Given a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $b \in \mathbb{R}^m$, the goal is to find a vector $x \in \mathbb{R}^n$ such that $Ax \approx b$. A common scenario in *deep learning* is having more parameters than inputs. i.e., $m < n$. Assume $m < n$, and loss function $\ell(x) = \|Ax - b\|^2$. Also assume that A is full rank. i.e., the rows of A are linearly independent.

   Suppose we run gradient descent starting at $\vec{0}$ for sufficiently many iterations, and we find a vector x^* such that $Ax^* = b$. Show that among all y such that $Ay = b$, the vector x^* has the smallest norm.

   (*Hint: Show that each of the vectors $x^{(t)}$ during gradient descent are in the row-span of A. Then show that if two vectors y_1, y_2 satisfy the equation $Ay = b$, then $y_1 - y_2$ is orthogonal to the row-span of A.*)

3. Let $\mathcal{H}_1$, and $\mathcal{H}_2$ be two hypothesis classes with functions from X to Y. Define $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2$ as a class of functions from X to $Y \times Y$ as follows: a function $f = (h_1, h_2) \in \mathcal{H}_1 \times \mathcal{H}_2 = \mathcal{H}$ is defined as $f(x) = (h_1(x), h_2(x))$. Show:

$$\tau_{\mathcal{H}}(m) \leq \tau_{\mathcal{H}_1}(m) \cdot \tau_{\mathcal{H}_2}(m)$$

4. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hypothesis classes with $\mathcal{H}_1$ containing functions from X to Y, and $\mathcal{H}_2$ having functions from Y to Z. Define $\mathcal{H} = \mathcal{H}_2 \circ \mathcal{H}_1$ as a class of functions from X to Z as follows: for all $h_1 \in \mathcal{H}_1$, $h_2 \in \mathcal{H}_2$, there exists a function $f \in \mathcal{H}$ defined as $f(x) = h_2(h_1(x))$. Show:

$$\tau_{\mathcal{H}}(m) \leq \tau_{\mathcal{H}_1}(m) \cdot \tau_{\mathcal{H}_2}(m)$$