AI534 — Written Homework Assignment 2 — Due Oct 29th, 2021

1. (Subgradient) (5 pts) Consider the L_1 form function for $\mathbf{x} \in R^d$: $f(\mathbf{x}) = |\mathbf{x}|_1 = \sum_{i=1}^{d} |x_i|$. Show that $\mathbf{g} = [g_1, g_2, ..., g_d]^T$ is a subgradient of $f(\mathbf{x})$ at $\mathbf{x} = \mathbf{0}$ if every $g_i \in [-1, 1]$. Hint: go back to the definition of subgradient: g is a subgradient of $f(x)$ at x_0 if $\forall x$, $f(x) \geq f(x_0) + g^T(x - x_0)$

2. (Perceptron) (5 pts) Consider the following argument. We know that the number of steps for the perceptron algorithm to converge for linearly separable data is bounded by $(\frac{D}{\gamma})^2$. If we multiple the input $\mathbf{x}$ by a small constant α, which effectively reduces the bound on $|\mathbf{x}|$ to $D' = \alpha D$, we can reduce the upper bound to $(\alpha \frac{D}{\gamma})^2$. Is this argument correct? Why?

3. (Cubic Kernels.) (10 pts) In class, we showed that the quadratic kernel $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^2$ was equivalent to mapping each $\mathbf{x} = (x_1, x_2) \in R^2$ into a higher dimensional space where

$$\Phi(\mathbf{x}) = (x_1^2, x_2^2, \sqrt{2}x_1x_2, \sqrt{2}x_1, \sqrt{2}x_2, 1).$$

   Now consider the cubic kernel $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^3$. What is the corresponding Φ function?

4. (Kernel or not). In the following problems, suppose that K, K_1 and K_2 are kernels with feature maps ϕ, ϕ_1 and ϕ_2. For the following functions $K'(x, z)$, state if they are kernels or not. If they are kernels, write down the corresponding feature map, in terms of ϕ, ϕ_1 and ϕ_2 and c, c_1, c_2. If they are not kernels, prove that they are not.

   - (5 pts) $K'(\mathbf{x}, \mathbf{z}) = cK(\mathbf{x}, \mathbf{z})$ for $c > 0$.
   - (5 pts) $K'(\mathbf{x}, \mathbf{z}) = cK(\mathbf{x}, \mathbf{z})$ for $c < 0$.
   - (5 pts) $K'(\mathbf{x}, \mathbf{z}) = c_1K_1(\mathbf{x}, \mathbf{z}) + c_2K_2(\mathbf{x}, \mathbf{z})$ for $c_1, c_2 > 0$.
   - (5 pts) $K'(\mathbf{x}, \mathbf{z}) = K_1(\mathbf{x}, \mathbf{z})K_2(\mathbf{x}, \mathbf{z})$.