CSE 151A Winter 2021 - Take Home Exam 5

1 hour **Name:** **PID:**

---

(1) This is an open book, open notes exam. You are free to consult any text book or notes. **You are not allowed to consult with any other person.**
(2) If you need any clarification, please post a private message to the instructors on Piazza.
(3) Remember that your work is graded on the *clarity* of your writing and explanation as well as the validity of what you write.
(4) This is a one-hour exam.

(1) State whether the following statements are true or false. Justify your answer.
   (a) (5 points) Any diagonal matrix is a valid kernel matrix. (Recall that a matrix A is called diagonal if for all $i \neq j$, $A_{ij} = 0$.)
   Solution
   False.
   Counter example: Consider the diagonal matrix $K = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. For any vector $v = [v1, v2]^T \in R^2$, $v^T K v = [v1, v2] \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} v1 \\ v2 \end{bmatrix} = -v1^2 - v2^2 \leq 0$. Therefore, there exists v (For example, $v = [1, 0]^T$) such that $v^T K v < 0$. Hence, K is not PSD $\implies$ K is not a kernal matrix.
   (b) (5 points) If $K(x, z)$ is a kernel function, then for all x and z, $K(x, x) + K(z, z) - 2K(x, z) \geq 0$.
   Solution
   True.
   Since $K(x, z)$ is a kernel, there exists a feature map ϕ such that $K(x, z) = \langle \phi(x), \phi(z) \rangle$.

$$\begin{aligned} K(x,x) + K(z,z) - 2K(x,z) &= \langle \phi(x), \phi(x) \rangle + \langle \phi(z), \phi(z) \rangle - 2\langle \phi(x), \phi(z) \rangle \\ &= \langle \phi(x) - \phi(z), \phi(x) - \phi(z) \rangle \\ &= ||\phi(x) - \phi(z)||^2 \\ &\geq 0 \end{aligned}$$

(2) For each of the following functions, state if they are a kernel function. If yes, provide the corresponding feature map; if no, provide a short proof.

(a) (5 points) S and T are subsets of $\{1, 2, \dots, 100\}$. $K(S, T) = |S \cup T|$.

Solution

False.

Let $x_1 = \phi$, an empty set and $x_2 = \{1\}$. Consider the matrix

$K = \begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) \\ K(x_1, x_2) & K(x_2, x_2) \end{bmatrix} = \begin{bmatrix} |\phi \cup \phi| & |\phi \cup 1| \\ |1 \cup \phi| & |1 \cup 1| \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$.

For $v = [1, -1]^T, v^T K v = -1 < 0$, which implies K is not a PSD matrix. Therefore, $|S \cup T|$ is not a kernel.

(b) (5 points) S and T are subsets of $\{1, 2, \dots, 100\}$. $K(S, T) = |S \cap T|$.

Solution

True. This kernel function is sinilar to the string kernel for the case where p=1 i.e it represents the number of common elements/letters in S and T. Therefore, we can define a similar feature map in this case as well.

Consider the feature map $\phi(x) = [\phi_1(x), \phi_2(x), \dots, \phi_{100}(x)]^\top$ where $\phi_k(S) = I(k \in S) = \begin{cases} 1 & k \in S \\ 0 & k \notin S \end{cases}$.

We can see that $K(S, T) = |S \cap T| = \sum_{i=1:100} \phi_i(S).\phi_i(T) = \langle \phi(S), \phi(T) \rangle$. For instance, let $S = \{1, 2\}$ and $T = \{2, 4\}$. $\phi(S) = [1, 1, 0, 0, ..., 0]$ and $\phi(T) = [0, 1, 0, 1, 0, 0, ..., 0]$. $K(S, T) = |S \cap T| = |\{2\}| = 1$. $\langle \phi(S), \phi(T) \rangle = \langle [1, 1, 0, 0, ..., 0], [0, 1, 0, 1, 0, 0, ..., 0] \rangle = 1.0 + 1.1 + 0.0 + 0.1 + 0.0 + 0.0 + ... + 0.0 = 1$. Thus, K is a kernel because there exists a feature map $\phi(.)$ such that $K(S, T) = \langle \phi(S), \phi(T) \rangle$.