# HW6

CISC648010 - Spring 2022

Due Date: April 8th at 11 PM

## 1 Kernel Functions 10 pts

a) To what feature map $\phi$ does the kernel $k(x,z) = x^T z + (x^T z)^2 + 4$ correspond? Assume that $x$ and $z$ have dimension $d = 3$.
b) Assume kernel function $k_1(x,z)$ corresponds to feature vector $\phi_1(x)$, and $k_2(x,z)$ corresponds to $\phi_2(x)$. To what feature map does kernel $k(x,z) = k_1(x,z) + k_2(x,z)$ correspond to (find the feature vector in terms of $\phi_1$ and $\phi_2$)? Remark: in general if $k_1(x,z)$ and $k_2(x,z)$ are valid kernel functions, then $k_1(x,z) + k_2(x,z)$ is a valid kernel too.

## 2 SMO algorithm for SVM 20 pts

Consider the following dataset:

$$x_1 = [1,0]^T, x_2 = [2,-1]^T, x_3 = [-1,1]^T$$

$$y_1 = 1, y_2 = 1, y_3 = -1$$

Consider the linear SVM problem where $\phi(x) = x$ and $k(x,z) = x^T z$.
a) Solve the primal optimization problem and find optimal $w$ and $b$. The primal problem is,

$$\min_{w,b} \frac{1}{2}||w||^2$$

$$s.t., y_i(w^T x_i + b) \geq 1, \forall i$$

b) Solve the dual optimization problem and find optimal $\alpha_1, \alpha_2, \alpha_3$. Using optimal values for $\alpha_1, \alpha_2, \alpha_3$ find optimal $w$. Note that the optimal $w$ found using $\alpha_1, \alpha_2, \alpha_3$ should be equal to the optimal $w$ found in part a. The dual optimization problem is,

$$\max_{\alpha_1,\alpha_2,\alpha_3} \sum_{i=1}^{3} \alpha_i - \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3} \alpha_i\alpha_j y_i y_j x_i^T x_j$$

$$s.t., \sum_{i=1}^{3} \alpha_i y_i = 0, \alpha_i \geq 0, \forall i$$

In order to solve the above optimization problem, use the SMO algorithm. Initialize $\alpha_1 = \alpha_2 = \alpha_3 = 0$ and run the SMO algorithm for one iteration. In this iteration, you only update $\alpha_1$ and $\alpha_3$. In this case, SMO finds the optimal solution only by one iteration. Do not use any calculators or any programming languages.

# 3 Nonlinear SVM 20pts

In this part, you are allowed to use the sklearn python library. The following code on googlecolab generates a dataset and determine the decision boundary using linear SVM (i.e., Linear Kernel $\phi(x) = x$ and $k(x, z) = x^T z$):

https://colab.research.google.com/drive/1wQUng7v3i9PxD8VDKkdtkW7Gx4jtNUwv?usp=sharing

You change the code to determine the decision boundaries using the following kernels:

- Polynomial Kernel with degree 3
- Polynomial Kernel with degree 10
- Gaussian Kernel with $\gamma = \frac{1}{2\sigma^2} = 0.1$
- Gaussian Kernel with $\gamma = \frac{1}{2\sigma^2} = 1$

Plot and save 4 figures that shows the decision boundary. Please include them in your report. Upload your code in a single file named SVM_lastname.py on canvas to get the full score.

Remark: Sklearn uses the term "Radial basis function (RBF) kernel" for Gaussian kernel. Using sklearn, you can set the parameter $\gamma$ not $\sigma$.

$$\text{Gaussian Kernel}: \; k(x, z) = \exp\{-\frac{||x-z||^2}{2\sigma^2}\} = \exp\{-\gamma||x-z||^2\}, \gamma := \frac{1}{2\sigma^2}$$