Kernel Methods
ELL409: Assignment-2

(Last updated: 08-24-2022, 2:00pm)

Deadline

Deadline for final submission of the assignment report for Assignment 2: **19 October 2022, 11:59 PM.**

Instructions

1. While you are free to discuss all aspects of the assignment with your classmates or others, your results and report must be entirely your own. We will be checking for any copying, and this will be treated as plagiarism and dealt with accordingly. In case of any doubts in this regard, please ask the TAs.

2. You should prepare a report compiling all your results and your interpretation of them as described in the problem statements, along with your overall conclusions. In particular, you should attempt to answer all of the questions posed in the respective parts of the assignment below. Any graphs or other visualizations suitable for description of the approach should also be included therein.

3. No hand-written solutions will be accepted or evaluated, and reports are expected to be submitted before the final deadline (as communicated in class). You can prefer reports in pdf format, however other formats might also be acceptable. In case of any doubts in the format of submission, please ask the TAs.

4. The schedule for demos/vivas will be announced by your respective TAs, in advance. If for any reason you cannot attend in your scheduled slot, you must arrange for an alternative slot with your TA well in advance. Last-minute requests for rescheduling will normally not be accepted.

5. The approach to the problems and results obtained must be reproducible. Failure to reproduce the results during demo(s) may result in penalty or considered as plagiarism, if suitable.

Part 1: Kernel Functions and Feature Maps (10 points)

The objective of this part of the assignment is to develop an understanding on concepts of kernel methods that you have learnt in the class. In this part, you will understand what kinds of kernel functions correspond to feature maps and develop correlations between them.

Kernel functions induced by feature map ϕ, are defined in the following manner:

$$K(x,z) = <\phi(x), \phi(z)>$$

Consider the kernel function, $K(\,.\,,\,.\,)$ defined as:

$$K(x,z) = (x^T z)^2$$

where $x,\ z \in \mathbb{R}^d$. Show that, $\exists$ feature mapping ϕ for which

$$K(x,z) = <\phi(x), \phi(z)>$$

Also, find the corresponding feature mapping ϕ.

The computational efficiency varies proportionally to the dimensionality of the mapping or function. Compute the computational efficiency (in terms of time complexity) for the following kernel functions and their feature mapping ϕ :

1. $K(x,z) = (x^T z)^2$
2. $K(x,z) = (x^T z + c)^2$
3. $K(x,z) = (x^T z)^k$

You are free to choose any notation for evaluation of time complexity (for example: Big-O). In a high-dimensional feature space, which amongst Kernel mapping or feature mapping would you preferred to use? Provide a justification for the conclusion based on results from the previous parts.

Part 2: Kernel Algebra (7 points)

The objective of this part of the assignment is to develop understanding of Kernel Algebra and formulation correspondence between kernel function and feature mapping. Kernel Algebra refers to the correspondence properties with feature maps. Deduce a general form of feature map composition for the following below kernel functions. An example has been provided below for reference.

1. $K\ (x, z) = \ f k_a(x, z),\ f > 0$
2. $K\ (x, z) = \ x^T A z,\ \ A$ is positive semi-definite
3. $K\ (x, z) = \ f(x)\, f(z)\, k_a(x, z)$
4. $K\ (x, z) = \ k_a(x, z)\, k_b(x, z)$

Example: For the kernel function given below:

$$K(x, z) = \ k_a(x, z) + \ k_b(x, z)$$

corresponding feature composition is of the form:

$$\phi(x) = (\phi_a(x), \phi_b(x))$$

Part 3: Valid Kernel Functions (8 points)

The objective of this part of the assignment is to understand validity of kernel functions and get introduced to Gaussian kernel function. By definition, $K(x, z)$ is a valid kernel **if and only if**

$$\exists\ \phi : X \to \mathbb{R}^d \ \text{ s.t}$$

$$K(x, z) = < \phi(x), \phi(z) >$$

The definition directly implies symmetricity of kernel functions (i.e., $K(x, z) = K\ (z, x)$).

Radial Basis Function or *Gaussian Kernel* for a bandwidth $\sigma > 0$, defined as:

$$K\ (x, z) = \ \exp\left(-\frac{\|x - z\|_2^2}{2\sigma^2}\right)$$

Note: The kernel function is based on Euclidean distance in n-dimensional space.

Based on the above definitions and properties,

1. Prove that Gaussian Kernel is a valid kernel.
2. Furthermore, show that

$$K(x, z) = \left(1 + \left(\frac{x}{\|x\|_2}\right)^T \left(\frac{z}{\|z\|_2}\right)\right)^3$$

   is a valid kernel.

   (**Hint:** $K(x, z) \ = \ \ x^T z$ is a kernel)