Assignment 4

This assignment is due **Monday November 23** at **11:59PM**.
Solutions should be turned in through the assignment FTP site in PDF form. The name of the PDF file should be ps4_YourStudentID, e.g., "ps4_9762578.pdf". Your MATLAB code (with comments) should be included in the PDF file too. The FTP site is 140.114.71.2 port 1235. You can use the same ID and password of the course website to login.

---

1. **(20 points)**

   Find $\boldsymbol{\phi}(\mathbf{x}) : \mathbb{R}^2 \mapsto \mathbb{R}^\infty$ for the 'Gaussian' kernels

   $$k(\mathbf{x}, \mathbf{z}) = \boldsymbol{\phi}(\mathbf{x})^\mathsf{T} \boldsymbol{\phi}(\mathbf{z}) = \exp\{-\gamma \|\mathbf{x} - \mathbf{z}\|^2\} ,$$

   where $\mathbf{x}, \mathbf{z} \in \mathbb{R}^2$, and $\gamma > 0$. You may assume $\gamma = 1$ for the sake of brevity.

2. **(20 points)**

   Consider the polynomial kernels of the form

   $$k(\mathbf{x}, \mathbf{z}) = (\mathbf{x}^\mathsf{T} \mathbf{z} + c)^n ,$$

   where $\mathbf{x}, \mathbf{z} \in \mathbb{R}^d$. What is the dimension of the feature space mapped by the implicit function $\boldsymbol{\phi}(\mathbf{x})$?

3. **(10 points)** 'Kernelizing' a neighbor-neighbor classifier. (PRML Exercise 6.3)

   The nearest-neighbor classifier assigns a new input vector $\mathbf{x}$ to the same class as that of the nearest input vector $\mathbf{x}_n$ from the training set, where in the simplest case, the distance is defined by the Euclidean metric $\|\mathbf{x} - \mathbf{x}_n\|^2$. By expressing this rule in terms of scalar products and then making use of kernel substitution, formulate the nearest-neighbor classifier for a general nonlinear kernel. (To begin with, you may assume an implicit mapping $\boldsymbol{\phi}(\mathbf{x})$ and substitute $\mathbf{x}$ by $\boldsymbol{\phi}(\mathbf{x})$. Make sure that the kernelized NN classifier only needs to evaluate kernels, and no explicit computation of $\boldsymbol{\phi}(\mathbf{x})$ is needed.)

4. **(20 points)** 'Kernelizing' Fisher's linear discriminant for two classes.

   Consider the Fisher criterion in the form

   $$J(\mathbf{w}) = \frac{\mathbf{w}^\mathsf{T} \mathbf{S}_\mathrm{B} \mathbf{w}}{\mathbf{w}^\mathsf{T} \mathbf{S}_\mathrm{W} \mathbf{w}} ,$$

as shown in Eq. (4.26) on page 189 of PRML. Change Eq. (4.20) into $y = \mathbf{w}^\mathsf{T}\boldsymbol{\phi}(\mathbf{x})$ and derive a 'kernelized' version of Fisher' linear discriminant for two classes. Note that the kernelized version involves only kernel evaluations. The implicit function $\boldsymbol{\phi}(\mathbf{x})$ should not appear in the final result.

5. **(30 points)** Gaussian Processes for Regression (MATLAB)

   In this problem you will use Eq. (6.68) on page 309 of PRML to do regression. The data set is 'space_ga.mat'. The training and test data are stored as matrices `data_train` and `data_test`. The target values are stored in `t_train` and `t_test`. Your program needs to predict the target values of the test data using Gaussian Processes with Gaussian kernels $k(\mathbf{x}, \mathbf{z}) = \exp\{-0.5\,\|\mathbf{x} - \mathbf{z}\|^2/\sigma^2\}$. Try to find suitable values for parameters β and σ^2, and report the error rates in terms of sum-of-squares error. Your code must be included in the PDF.