You may submit the homework in pairs. Write both full names when submitting.

1. In class we assumed that $\mathbb{E}[\phi(x)] = 0$. If this is not the case, the first step is to apply centering before taking the eigendecomposition of K. Show that this is equivalent to computing the eigendecomposition of,

$$K_c = K - \mathbf{1}_N K - K\mathbf{1}_N + \mathbf{1}_N K \mathbf{1}_N$$

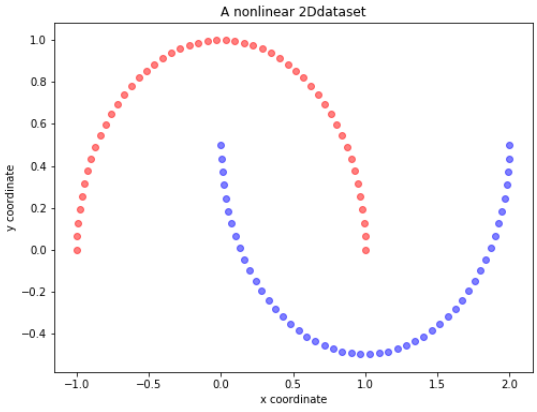
   where 1_N is a matrix of $1/N$ in all elements.

2. Based on our derivation of the Nyström extension in class, prove that under the low rank assumption, we can compute K_{22} via,

$$K_{22} = K_{21} K_{11}^{-1} K_{12}$$

3. Coding problem: For the two moons dataset and for the circles dataset, find a parameter setting for which kernel PCA successfully separates the data into two clusters and a parameter setting for which it does not. Explain in 2-3 sentences why it fails.

   - For two moons, in Python use sklearn.datasets.make_moons with 100 samples or generate the data yourself: upper half of a circle with radius 1 whose center is at (0,0) and lower half of a circle with radius 1 whose center is at (1,0.5).



   - For the circles dataset, generate 200 randomly sampled points along a circle of radius 1 and 200 randomly sampled points along a circle of radius 0.25. Add noise of std=0.1 to the locations of the points.

4. Coding problem: Generate 2000 points from the two moons dataset.

   - Randomly sample 500 points and perform kernel PCA for the full dataset using the Nyström extension.
   - Randomly sample 500 points only from the left moon and perform kernel PCA for the full dataset using the Nyström extension.

   Plot the first components for both results and explain.