(1) Fix $n > 0$. For $x, y \in \{1, 2, \ldots, n\}$ define $k(x, y) = \min(x, y)$. Give an explicit feature map $\varphi : \{1, 2, \ldots, n\}$ to $\mathbb{R}^D$ (for some D) such that $k(x, y) = \varphi(x)^T \varphi(y)$. Hint: construct $\varphi : \mathbb{R} \to \mathbb{R}^n$ using indicator functions $\mathbf{1}(x \leqslant i)$.

(2) Consider the objective function

$$J(w) = \|Xw - y\|_1 + \lambda \|w\|_2^2.$$

Assume we have a positive semidefinite kernel k.
   (a) What is the kernelized version of this objective?
   (b) Given a new test point x, find the predicted value.

(3) Suppose you are given an training set of distinct points $x_1, x_2, \ldots, x_n \in \mathbb{R}^n$ and labels $y_1, \ldots, y_n \in \{-1, +1\}$. How can you select σ to achieve perfect $0 - 1$ loss on the training data using a linear decision function and the RBF kernel.

(4) Suppose you are performing standard ridge regression, which you have kernelized using the RBF kernel. Prove that any decision function $f_\alpha(x)$ learned on a training set must satisfy $f_\alpha(x) \to 0$ as $\|x\|_2 \to \infty$.

(5) Complete the missing functions in *hw4.ipynb* to implement kernelized ridge regression with the RBF kernel. Include the generated plots showing the actual function and the prediction. Can you interpret the plots using the solution of the previous problem.