CX 4803 CML – Spring 2022

Assignment 7

1. In ridge regression, a prediction $\hat{y}$ is made for a value of x via the linear model $\hat{y} = w^T x$, where w is a vector of weights. The weights are found via

$$\min_w \left[\frac{1}{2} \sum_i (y_i - w^T x_i) + \frac{\lambda}{2} w^T w \right]$$

   where λ is a regularization parameter and we are given data pairs (x_i, y_i).

   Now we wish to "kernelize" ridge regression so that we can fit the data with a nonlinear function. Let the kernel function be denoted as $k(\cdot, \cdot)$. Explain how a prediction $\hat{y}$ is computed for this kernelized method.

2. The code `demo_svm.m` for a support vector classifier is provided on Canvas (in the Apr-04 lecture notes folder).

   (a) First convert this code to use a Gaussian kernel function

$$\exp(-\frac{1}{2s^2} \|x - x'\|_2^2)$$

   where s is the "standard deviation" of the kernel function.

   (b) Generate an *interesting* data set in 2D that you can use to test the modified code, which is now a support vector machine (SVM). Your data set should have overlapping samples from two classes and the two classes should ideally be separated by a nonlinear discriminant. (The synthetic data set given in `demo_svm.m` is not interesting because the classes are ideally separated by a straight line.) Your interesting data set could be a manually selected set of points.

   (c) For your data set, test SVM using different values of s in the kernel function as well as different values of the parameter C used in the SVM dual objective function. Plot your results, showing the data set (labeled differently for each class) and the discriminant. Explain why the results are expected or not expected.