EN3150 Assignment 04: Kernel Methods

Sampath K. Perera

1 Kernel Methods

1. Suppose that input space to feature space mapping (projection) is given by the following function.
$$\Phi(x) = (1,\ \sqrt{2}x\ , x^2). \tag{1}$$
For a one-dimensional input space show that the corresponding kernel function is $k(x,z) = (1 + xz)^2$.

2. Express the kernel function provided above for a two-dimensional input space, where $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{z} = (z_1, z_2)$.

3. What is the mapping function $\Phi(\boldsymbol{x})$ for the kernel function provided above in the above question (Q2).

4. Consider the kernel $k = (1 + \boldsymbol{x}^T \boldsymbol{z})^2$. For the following dataset, determine the kernel matrix (gram matrix). Is this a valid kernel for this data set? justify your answer.

Sample index	Data sample ($\boldsymbol{x} = (x_1, x_2)$)	Feature 1 (x_1)	Feature 2 (x_2)
1	$\boldsymbol{x}_1$	1	5
2	$\boldsymbol{x}_2$	3	4
3	$\boldsymbol{x}_3$	4	2
4	$\boldsymbol{x}_4$	10	12

Note:- The kernel matrix (gram matrix) is given by

$$\boldsymbol{G} = \begin{bmatrix} k(\boldsymbol{x}_1, \boldsymbol{x}_1) & k(\boldsymbol{x}_1, \boldsymbol{x}_2) & \cdots & k(\boldsymbol{x}_1, \boldsymbol{x}_N) \\ k(\boldsymbol{x}_2, \boldsymbol{x}_1) & k(\boldsymbol{x}_2, \boldsymbol{x}_2) & \cdots & k(\boldsymbol{x}_2, \boldsymbol{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(\boldsymbol{x}_N, \boldsymbol{x}_1) & k(\boldsymbol{x}_N, \boldsymbol{x}_2) & \cdots & k(\boldsymbol{x}_N, \boldsymbol{x}_N) \end{bmatrix} \tag{2}$$

5. Use the code given in listing 1 to generate data.

```python
import numpy as np
import matplotlib.pyplot as plt
from sklearn.datasets import make_circles
# Generate data with make_circles
np.random.seed(5)
X, y = make_circles(n_samples=500, factor=0.3, noise=0.1)
```

Listing 1: Data generation.

   (a) Plot the scatter plot of the data (plt.scatter (X [:, 0], X [:, 1], c=y, cmap=plt.cm.Paired)).
   (b) Use the following mapping to map two-dimensional space to three-dimensional space (feature space). This feature space set is known as projected set. Visualize the projected set.

$$\Phi : \boldsymbol{x} = (x_1,\ x_2) \to \Phi(\boldsymbol{x}) = \left(x_1,\ x_2,\ x_1^2 + x_2^2\right) \in \mathbb{R}^3.$$

   (c) Now change the "factor=0.5" and observe the visualize the feature space in 3-d. What changes can you observe?
   For the same data, use this mapping: $\boldsymbol{x} = (x_1,\ x_2) \to \left(x_1^2,\ x_2^2,\ x_1x_2\right)$. Is this mapping better than the previous mapping?
   (d) Run linear SVC (svm.SVC(kernel='linear')) on the original dataset which is generated based on listing 1 and the projected set using the mapping given in 5b and report the classification accuracies for both cases.

2 Additional Resources

1. Scikit-learn SVM documentation
2. Scikit-learn Custom Kernel
3. Stanford CS229 Lecture notes

Submission

- Upload a report and your codes as a zip file named as "EN3150_your_indexno_A04.zip". Include the index number and the name within the report as well.
- The interpretation of results and the discussion are important in the report.
- An extra penalty of 10% is applied for late submission.
- Plagiarism will be checked and in cases of plagiarism, an extra penalty of 10% will be applied.