## SVM Problems Solutions:

1. Show that a linear separator can always be found in this embedded space, regardless of radius and where the data is centered

Proof:

It is easy to see that for the original data, consider data that is distributed such that anything within distance ***r*** of the center ***(a, b)*** is of class +1, and anything greater than distance ***r*** of the center is of class -1, there is no linear separator, but the data can be classified correctly with . And it is easy to find that,

So, it is obviously that we have a linear separator for the feature map  with,

So, regardless of radius ***r*** and center ***(a, b)***, there is always a linear separator in this embedded space.

1. show that if there is an ellipsoidal separator, regardless of center, width, orientation a separator can be found in the quadratic feature space using this kernel.

Proof:

As the same reason in the **1**, and we already know the function of ellipsoidal in high dimensional coordinate system, then we have the function of the separator for the original data set:

Where is the center, n is the dimension.

So it is easy to find that we have a linear separator for the feature map with,

So it is easy to see that there is a separator in the quadratic feature space .

Suppose we have two ellipsoidal regions denoted as **A** and **B**, from the proof of question 1, we can have two separate formula to denote the two-dimensional data set that satisfied the following: the positive class lay within one ellipsoidal region and the negative class was everywhere else. So we can write the following formula:

For **A**: the data within **A** is positive class else negative,

For **B**: the data within **B** is positive class else negative,

So we can conclude that we have three situations that satisfied our question.

1. The data point is within **A** and it is not in **B** so we have .
2. The data point is not in **A** and it is within **B** so we have .
3. The data point is not in **A** and it is not in **B** so we have .

So we can have the following relationship to describe the condition mentioned in the question:

If , data point is either in **A** or in **B**, else .

So,

Therefore, because it is easy to show that can be a linear separator for feature map , we can know that the kernel will recover a separator for this question.

Suppose we have two circle regions denoted as **A** (inner circle with radius ***r*** and center ()) and **B** (outer circle which contains **A** with radius ***R*** and center ()), from the proof of question 1, we can have two separate formula to denote the two-dimensional data set that satisfied the following: the positive class lay within **A** and the negative class within circular band between **A** and **B** and positive everywhere else. So we can write the following formula separately:

Just for **A**: the data within **A** is positive class else negative,

For **B**: the data within **B** is positive class else negative,

So we can conclude that we have three situations that satisfied our question.

1. The data point is within **A** and **B** so we have .
2. The data point is in the band which means it is not in **A** and it is within **B** so we have .
3. The data point is not in **A** and it is not in **B** so we have .

So we can have the following relationship to describe the condition mentioned in the question:

If , data point is either in **A** and **B** or not in **A** and not in **B**, else .

So,

Therefore, because it is easy to show that can be a linear separator for feature map , we can know that the kernel will recover a separator for this question.

We already know that for XOR problems,

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

So we can assume and use to replace in case of confusion and from the lecture note we already know that,

The objective function which we need to maximize for the dual problem becomes

Where

with constraints and .

and we firstly calculate the :

The same way we get , .

So,

So,

Optimizing with respect to the Lagrange multipliers yields the following set of simultaneous equations:

So we get

The function of separator or boundary is which is:

So we can assume and use to replace in case of confusion

We can expand the kernel by using series , so

So feature map is:

The objective function which we need to maximize for the dual problem becomes

Where

with constraints and .

and we firstly calculate the :

The same way we get , , .

So,

So, we can put it into and optimizing with respect to the Lagrange multipliers yields the following set of simultaneous equations:

So we get

So the separator is

so the bounder is

So, we can defined the preference by using the margin

For Polynomial Kernel:

Margin is

For gaussian kernel, we have, .

But I don’t know how to solve .

But I assume that the polynomial kernel is better.