Theodore Kim

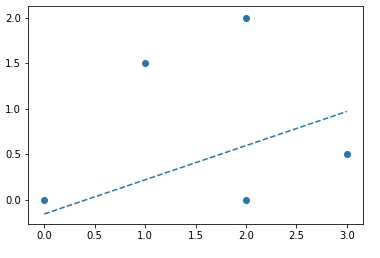
CS-UY 4563 – Introduction to Machine Learning

HW #5

1. Consider the training data:

|  |  |  |
| --- | --- | --- |
| **x1** | **x2** | **y** |
| 0 | 0 | -1 |
| 2 | 2 | -1 |
| 2 | 0 | 1 |
| 1 | 1.5 | -1 |
| 3 | 0.5 | 1 |

(a) Plotted data with hyperplane: **w** = [12, -32], **w0** = -5



(b) The data **is** linearly separable

(c) Given the above values

**i. Determine the functional margin:**

**ii. Determine the geometric margin:**

**iii. compute the canonical weights with respect to the training example**

The canonical weights with respect of the training set is the value for **w** and **w0** such that the functional margin is 1.

Therefore, the functional margin was previous **5** therefore divide w by 5 to get the canonical weights:

and

(d) Identify which of the training examples are support vectors

The training samples closest to the hyperplane are the support vectors, therefore the support vectors in this example are: (0, 0) on the negative side and (3, 0.5) of the positive side.

(e) Adding the point x = (1, 3) and y = -1 **does not** change the margin, hyperplane or support vector as it lies outside of the maximum margin as its functional margin is not equal to -1.

(f) Removing the point x = (1, 1.5) **does not** change the margin, hyperplane or support vector as the point previously lay outside of the maximum margin as its functional margin was not equal to -1.

(g) Removing the point x = (0, 0) **does** change the support vector as it then moves it outwards to the next closest point: (1, 1.5). In response the maximum margin increases (becomes wider) as both support vectors are now farther apart. Finally the separating hyperplane shifts in response to the wider maximum margin.

(h) Define a constraint optimization problem:

Subject to

2. Consider the training data:

|  |  |
| --- | --- |
| **x1** | **y** |
| 1 | -1 |
| 2 | 1 |
| 3 | 1 |
| 4 | -1 |

Create new features using the Gaussian Radial Basis Function, let σ = 1

For all values

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **x1** | **y** |  |  |  |  |
| 1 | -1 | 1.00000 | 0.60653 | 0.13534 | 0.01111 |
| 2 | 1 | 0.60653 | 1.00000 | 0.60653 | 0.13534 |
| 3 | 1 | 0.13534 | 0.60653 | 1.00000 | 0.60653 |
| 4 | -1 | 0.01111 | 0.13534 | 0.60653 | 1.00000 |

The coefficients calculated using hard margin SVM on the above features:

3. Running the hard margin SVM on the added data below using the weights would get the predictions:

For all values , z has been calculated using a radial basis function kernel

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **x1** |  |  |  |  |  |
| 1.2 | 1 | 0.429557 | 0.122303 | 0.245904 | -1 |
| 2.5 | 0.429557 | 1 | 0.75484 | 0.932102 | 1 |
| 3.25 | 0.122303 | 0.75484 | 1 | 0.932102 | 1 |

3. Would the two constrained optimizations create the same decision boundary?

**The second one would result in the same decision boundary, while the first one won't because, rather than creating support vectors of functional margins of one, it is creating a line which achieves the maximum marginal SVM.**