# HW3 – Gal Kaptsenel 209404409

## Q1

### 1.1

Let points in .

There is a subset of at most 4 points, each of the points in has an extreme value in at least one of the x/y axes (if there are two points with the same max/min value, take one of them). The size of is at most 4 because there are two axes (x and y), and for each of them we will take a point with maximum value and (maybe another) point with minimum value.

Any rectangle that contains must contain all the points in because given a point ,

Because contains the points from with the maximum and minimum coordinates in both axes.

is at most with size 4, therefore there is , given the labeling that labels with **false** and the rest of the points with **true**.

Suppose there exists a rectangle that contains all the points with the **true** label and does not contain the single point () with the **false** label. This rectangle contains all points in the set , and because , the rectangle contains all the points in . Therefore, this rectangle contains all points in , and explicitly also , which contradicts the assumption that the rectangle does not contain . Thus, the assumption is incorrect and there is no such rectangle.

Therefore, cannot shatter , and because is subset of any 5 points in , we can conclude that .

### 1.2

Given two hypothesis classes such that . Let be a set with size of that is shattered by , exists such set by the definition of .

Given labeling for set , there exists a hypothesis which completely agrees

with the labeling, and therefore shatters .

Therefore, , because there exists a set with size

which shattered by

### 1.3

It could be said that .

Given , we could construct a decision tree that implements (i.e. makes the exact same prediction as h).

The tree, which has a depth of 4, will make the following decision for a given ,

Check whether the coordinate is between and , if no, return -1 (2 comparisons, therefore this check requires a depth of 2), afterwards, check whether the coordinate is between and , if no return -1, otherwise return 1 (2 comparisons, therefore requires additional depth of 2). The total depth of this tree will be 4, .

For any given point for prediction, the described tree will return a prediction of 1, iff and , otherwise it will return , and therefore it makes the exact same predictions as .

Therefore, according to 1.1 and 1.2 above, .

Visualization of :

Check if

Check if

Check if

Check if

+1

-1

-1

-1

-1

Depth of 4

## Q2

Let ,

Therefore,

Moreover, it could be seen that is valid, according to the kernel algebra (lecture SVM slide 46):

it is given that are valid kernels, therefore, according to rule 3 form kernel algebra, and are valid kernels as well, and according to rule 4 from kernel algebra,

is a valid kernel.

## Q3

### 3.1



And therefore,

and therefore,

1. are convex functions.

In conclusion, by definition, is convex , because

### 3.2

According to the rule from Tutorial 01, that states that any linear function is convex (slide 13, lemma 1), we can conclude that both the functions



Are convex .

According to above, is convex

## 3.3

According to slide 12 from tutorial 7, is convex, and according to the given two lemma 1 in this question, and because , is convex.

According to given lemma 2 in this question and above, is convex.

According to lemma 1 in this question, is convex.

According to lemma 2 in this question, is convex.

In addition, is convex set because,

, because is a vector space.

Therefore, according to the property from slide 15, tutorial 7, which states that if we restrict a convex function to a convex subset, then it is a convex function,

the above , is convex.

We can conclude that is convex and therefore the problem ,

i.e. Soft-SVM, is a convex optimization problem.