# Optimization

## Assignment 2

Yonatan Simson, 015764921

Amit Kristal, 025602624

### Q1

In the above figure we wish find d, the minimal distance between point and line .

We do so by projecting onto , using the principal of orthogonality:

QED

### Q2

What makes this problem convex is the minimization of L2 norm in

In fact this is equivalent to the quadratic programming problem as learned in class:

Prove: is convex

Proof:

Due to triangle inequality of the L2 Norm

QED

Prove: is convex

Proof:

Ask Alex how a linear expression can be defined as convex

### Q3

Condition (4a) is constraint given in the Primal form.

To prove the rest we use KKT conditions. The Lagrangian for this problem is:

Differentiating (\*) by and will give us

We of course have from the Lagrangian

Proof of (4c)

Say we have the following Lagrangian of a general convex problem ()

Then the first order optimality condition is

If the constraint is inactive then

If the constraint is active then

Either way the equality holds true.

Using this we get:

The active constraints correspond to the case of

Where as the inactive constraints correspond to the case of

So we conclude that only the vectors corresponding to the active constrains play a role in defining . These are *the* support vectors that give SVM it's name.

### Q4

The Dual form is derived by taking the results from the previous question (4d), (4e) and substituting them into the Lagrangian:

Using strong duality we derive the dual form from the above equation:

QED

### Q5

Equation (4d) can be expressed in matrix form as:

So extracting ***w*** from is all we need to do.

In order to extract we make use of the inequality constraints when they are active.

We can also average over the entire active set to get a better estimate

### Q6

…

### Q7

….

### Q8

For the following box set

The projection of x on

### Q9

Matlab

### Q10

For the following problem:

The Lagrangian is

Deriving the KKT optimality conditions

If both are zero then this means both constraints are non active. This means that:

If only and then is an active constraint:

If only and then is an active constraint:

### Q11

For the following problem

I will approach this problem in this way. If I knew the solution I would know which of the constraints are active the problem would be easy to solve. But I do know what the solution to the unconstrained problem would be. . Using this can expose which constraints would be active . The since it is inactive.

Now we can use the Lagrangian to solve this problem analytically.

The final answer is:

When starting at the gradient at this point is:

The Hessian is: since we can write our function as

The since this is the required step size to get to from using

According to this . This just so happens to be the correct solution for our problem. The reason for this would be that the Hessian is Unitary matrix.

In order to make this point we can take a different example. Such as:

Where

The correct solution is found by using the projected Newton algorithm as described in the question 11. If however we use without using the reduced Hessian we get:

### Q12

The equation for the augmented Lagrangian for problem (10) is:

Where