Advanced topics in Machine Learning

Assignment 1

Mariuta Nicolae (rqt629) September 13, 2015

# Question 1

I assign random variable with sample space where value 0 means that passenger did not show up and 1 means that the passenger has shown up.

Since 5% of the 10000 passengers did not show up, it means that 9500 have shown up so we can estimate the expected value for 100 passengers:

By applying the Hoeffding’s inequality, I obtain the bound:

Because we want to calculate the bound the probability that the number of people that show up is larger than the number of seats, we have that

As a comparison, if I use Markov’s inequality to bound the same probability, I obtain:

# Question 2

For this problem, the objective function is the Euclidean distance between (x,y) and the point (-2,2):

And after doing some basic calculation, I could bring that to the form:

So that will be the function we want to minimize, but there is the constraint that the point (x,y) to be inside the circle with radius 1 around the origin, so we also have the constraint function :

So given this optimization problem, I define the Lagrangian function by combining the objective and the constraint function:

Now, I calculate 3 partial derivatives of with respect to , and .

Since it is a quite tricky function to calculate derivative and to avoid doing mistakes, I used an online tool to calculate the derivatives with respect to x and y. I obtained:

For solving the optimization problem, I obtain a system with 3 equations to be solved:

The optimal solution is found by solving the above system of equation, but I unfortunately did not manage to solve that one.

I am not sure how correct is, but I think I can approach the problem in a simpler way so I can make the calculations easier. Since we want to minimize the objective function:

I am also going to minimize the result of what is below the square root, so I will reduce to the function:

that is also going to minimize the Euclidean distance.

By using the same constraint function as previously, the Lagrangian function becomes:

And now it is a lot easier to calculate the derivatives:

With that simplification, I also get a system of equations that is a lot easier to solve:

Now, I can calculate the value of :

In the next step, I calculate the possible values of :

By making the replacements, I obtain the possible values for , x and y I obtain:

If I look at the graphic with the circle and the second pair of values for x and y( ) I see these are exactly the coordinates of the closest point on the circle to our objective (-2,2). The other pair of coordinates that I obtained are symmetric on the other side of y axis.

# Question 3

**1. Learning with**

For deriving this high probability bound, because is finite, I use the bound for the finite hypothesis classes:

Where is the hypothesis that takes strings from and returns {0,1} according to the prediction that is made. is the expected loss of and is the empirical loss of on the training set that is provided. is the size of and n is the number of samples from the training set. is our confidence budget.

With probability greater than we have:

So we can say that:

Because is the set of all functions from to {0,1} then I guess the size of is equal with the number of strings with size d from the set which will make that M to have a large value and it will result in a tight bound.

**2. Learning with**

Because in this case, the size of hypothesis space is infinite, we will have the bound:

For the high probability bound we have:

Where we have that

In this case, the bound is a lot wider because the budget is distributed along all the spaces with strings of different sizes.