**Machine Learning Exercise**

**Ruoxin WANG (2030026150)**

**2023/04/18**

**Q1. What linear function is used by a SVM for classification? How is an**

**input vector 𝐱𝐢(instance) assigned to the positive or negative class?**

(1) linear classifer:

(2) for each , when , we will assign assigned to positive class; when , we will assign assigned to positive class.

**Q2. If the training examples are linearly separable, how many decision boundaries can separate positive from negative data points? Which decision boundary does the SVM algorithm calculate? Why?**

(1) if training examples are linearly separable, we can take numerous decision boundaries, because there are infinite lines we can used to separate some linearly organized points.

(2) here SVM focus on the maximum margin to separate hyperplane. Because this solution can guarantee the best generalization performance. And we can make correct and confident predictions on data as much as possible.

**Q3. Use Lagrange multiplier method to answer the following questions.**

**3.1 Consider the Entropy definition:**

**– If we are given a probability distribution , then the information conveyed by this distribution, also called the Entropy of 𝑃,is (The base of the logarithm is 2)**

**What is the range of entropy of 𝑃 and which distribution gives maximum entropy? Show the details of your answer.**

In order to calculate the range, we can calculate lower bound and upper bound of entropy.

A. Since is a ratio which range between 0 and 1. So when one of , according to formula , we can conclude that other . And we have , which reach the minimum.

B. here let us calculate upper bound.

(1) we want to maximize the entropy function:

(2) then we can construct Lagrange multiplier:

(3) And we can calculate partial derivative of and respectively:

Therefore, is a constant:

(4) So, the upper bound is:

Finally we can conclude that the range of entropy of P is

**3.2 Given the probability distribution , Gini index is another way to measure the uncertainty** **. What is the range of Gini index and which distribution gives maximum Gini index value? Show the details of your answer.**

In order to calculate the range, we can calculate lower bound and upper bound of gini index.

A. lower bound.

(1) here we want to minimum the gini index of P:

(2) Since we know is a ratio which range between 0 and 1, so:

(3) When we want to minimize , we can let as large as possible, so here we replace by :

(4) Since we know:

B. Upper bound.

(1) here we want to maximize the gini index of P:

(2) then we can construct Lagrange multiplier:

(3) And we can calculate partial derivative of and respectively:

Therefore, is a constant:

(4) So, the upper bound is:

Finally, we can conclude that the range of entropy of P is

**Q4. Given the SVM optimization problem:**

**Derive the dual optimization problem and show the detail steps.**

(1) we can write the constraints as:

(2) construct the Lagrangian for optimization problem:

(3) we will do by setting the derivatives of with respect to w and b to 0:

(4) plug back into Lagrangian equation:

(5) then, we obtain the dual optimization problem:

**Q5. For a solution**  **for the dual problem, if exist that , then the solution for primal problem is:**

**Show the detail steps for deriving the solution given .**

(1) as for , we can derive the solution by taking derivative of Lagrangian formula:

When , we have:

(2) as for , we can solve the equation by plug in the definition of support vector.

Since the value of can only be either 1 or -1, so we rewrite equation as:

Since we have already find the value of w, plug in the above equation: