Intro To ML – HW4

students:

Dor Bank - 301380416 - [dorbank@gmail.com](mailto:dorbank@gmail.com),

Avi Caciularu - 203056585 - <avi.c33@gmail.com>

**Theory Questions**

**Question 1:**

1. Let be hypotheses and let be their majority:

.

Assume that for every feature , there are exactly hypotheses which return the correct label of , meaning: .

Then for every , the majority function will return the incorrect label, meaning .

Let be the hypotheses: . Explicitly:

Let we have 3 samples:

. Thus thus . We get that but (where ).

1. Assume now that . Then . In total, we get that .

**Question 2:**

1. The error of the current hypothesis relative to the new hypothesis can be expressed by:

where .

Thus, we will now show that :

1. Proof by contradiction that . Then:

.

But where is a class of weak learners, by the definition of AdaBoost. Hence should fulfill- . Thus,

**Question 3:**

1. Denote and *.*

Denote

Hence: *.*

Denote the matrix which all its elements are equal to as . Thus, generalizing the discussion to matrices, we get: and since and are symmetrical (by definition of ), then and we get: **.** Assuming we are given and that matrix multiplication takes for matrixes dimension’s , we conclude that the time complexity to compute in total is .

1. We will show that any where can be expressed by a linear combination of :

Denote as the covariance matrix formed of , meaning: . We note that every is an eigen-vector of , with eigen-value (by the definition of principal component). Thus:

.

Hence:

(note that we didn’t prove that , which is easy to show). Denote and we get: as we expected.

1. . Calculating and depends on the number of features, meaning it takes . Repeating this times (as the number of samples), and in total, this phase’s time complexity is .

**Question 4:**

Consider the following optimization problem:

Where is a matrix of the variables, is a column vector of size , and is a column vector of size of coefficients.

Consider the SVD of , we get the constraint:

We want now to recover , so from the recitation: . By multiplying both members with from left, we get: . Denote this solution by .

Assume is singular. Denote by a vector in the null space of . Namely a vector such that . Suppose other solution to the optimization problem: , so fulfills . Thus, we get: .

Hence any solution of the form contains all possible solutions to .

From the recitation, we know that the projection matrix, can be calculated by . We would like to project to the null space of so we can get the minimum distance vector from . Then the null space projection would be: , and for we get: .

Then, from the definition of projection, the closest vector to in this null space is **.**

We know that the null space is linear, hence as well.

Now we can solve the optimization problem:

.

**Programming Questions**

**The code source files stay under /specific/a/home/cc/students/cs/avicaciularu/ML/HW4/:**

**Q5- Q5.py**

**Q6- Q6.py**

**Question 5:**

**Question 6:**