**Machine Learning**

**Billio Jeverson | 1002939**

**Homework 3**

This is the report and instruction as to how to access my folder.

For each of the question, there is a folder allocated for it, all the extra materials all are attached in the folder.

**Qn1.)**

Kernel type 0 = Accuracy = 84.375% (27/32) (classification)

Kernel type 1 = Accuracy = 81.25% (26/32) (classification)

Kernel type 2 = Accuracy = 90.625% (29/32) (classification)

Kernel type 3 = Accuracy = 43.75% (14/32) (classification)

Inside this folder, above result is contained in “Accuracies.txt”.

The model for each of the kernel type is put under “training[kernel type].txt.model”

The output is saved in “output[kernel type].txt”

I.e. for Kernel type 0, model = training0.txt.model

output = output0.txt

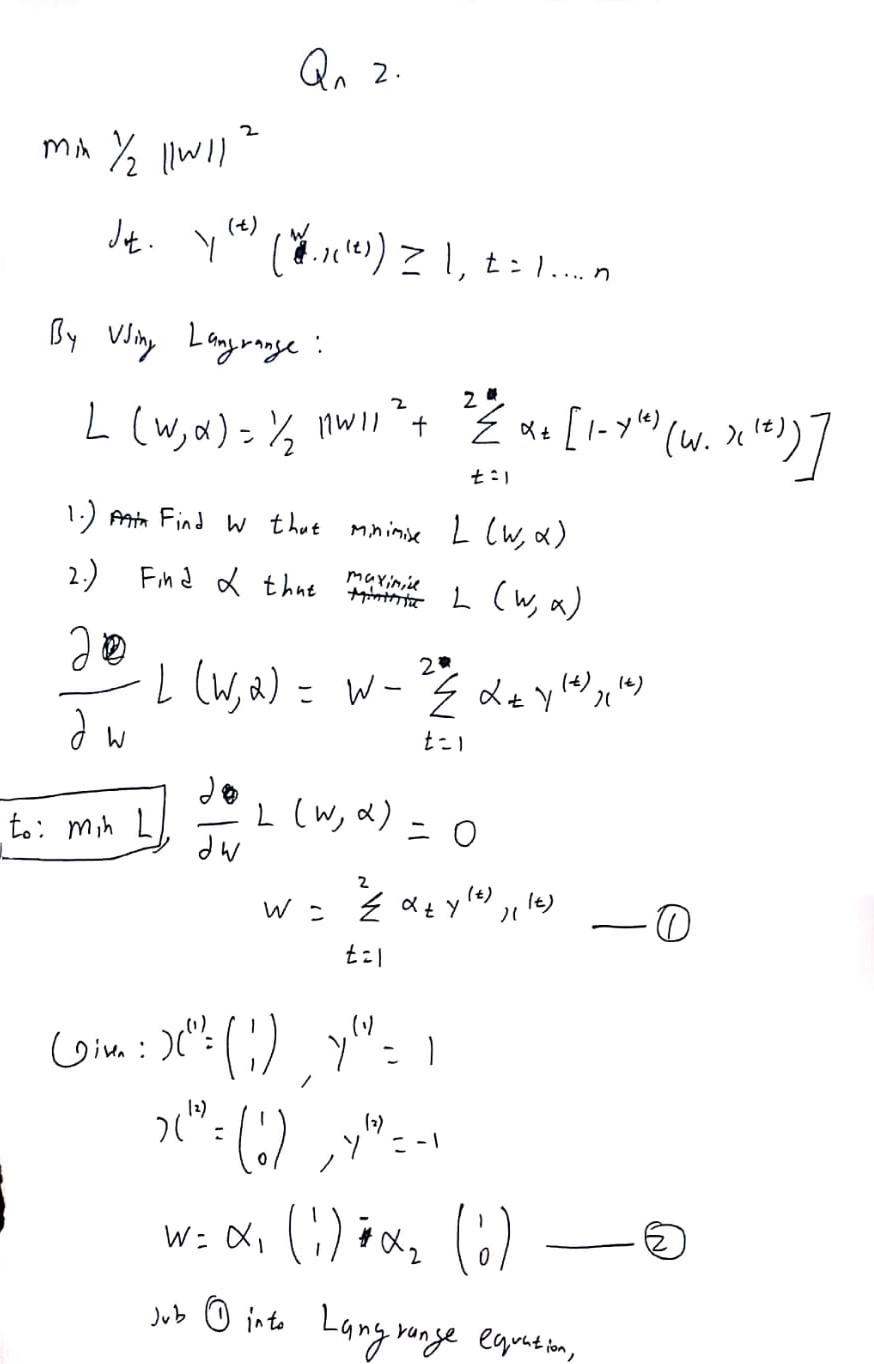
**Qn2a.)**

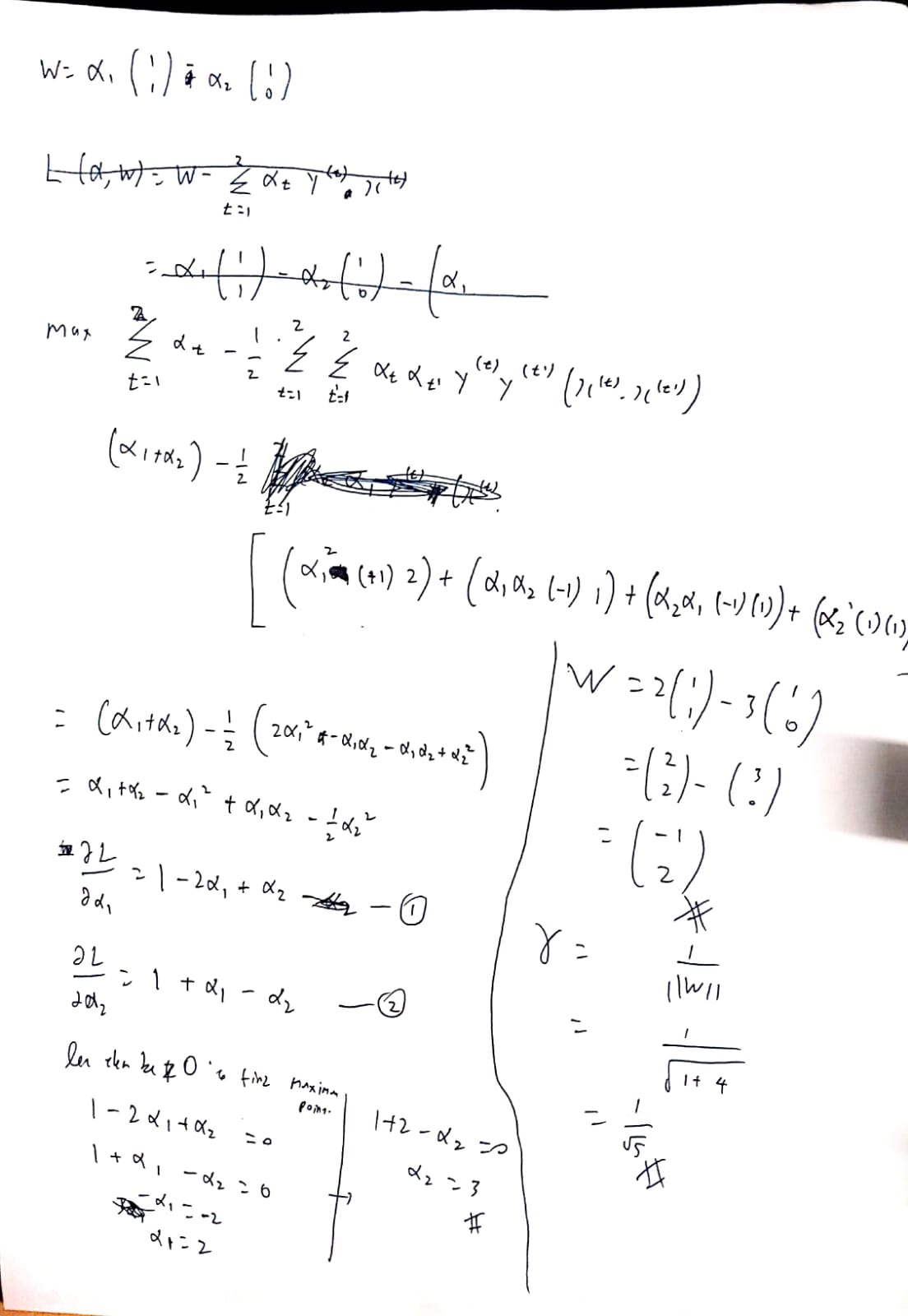
So in this question, I tried several methods but I get all to be the same answer so I am just going to put Langrange method here. Since it is the most confirm correct ones.

So, the idea is to find W to minimize a new function we have create L(a,w) , our langrange multiplier function.

Next is to find a to maximize the L(a,w).

In the first step, we will have w in terms of a then we sub it back to L(a,w) so that now L will only have a as its variable. Then we differentiate it and equatute it to 0. a is a vector this I have a1 and a2 and needed to be differentiated.

To ensure that a1 and a2 are indeed maximum points of L(a,w). I did not write it below but I differentiat L(a,w) w.r.t to a1 and a2 to a second degree. Both of them indeed obtain a -ve number of -2 and -1 for a1 and a2 respectively. This L is maximize and a1 and a2 given below. ****



**2b.)**

For this question, I will write my answer here. So we realise that the margin is 2/||W||. That being said since we only have 2 data points, the 2 data points will be the supporting vectors. Since the margin is also the distance between the 2 supporting vectors, the margin will not change even when set off b, is given.

The w however will change because now the SVM Hyperplane will be able to move “freeer” it will adjust itself to give the most accurate decision boundary. Separating the two data points.

Having written that, W vector does change but ||W|| does not change so as the margin will stay the same.

**Qn3.)**

In this question, I started with definition of Kernel and positive semi-definite matrix.

For part 1.) Kenel definition is used

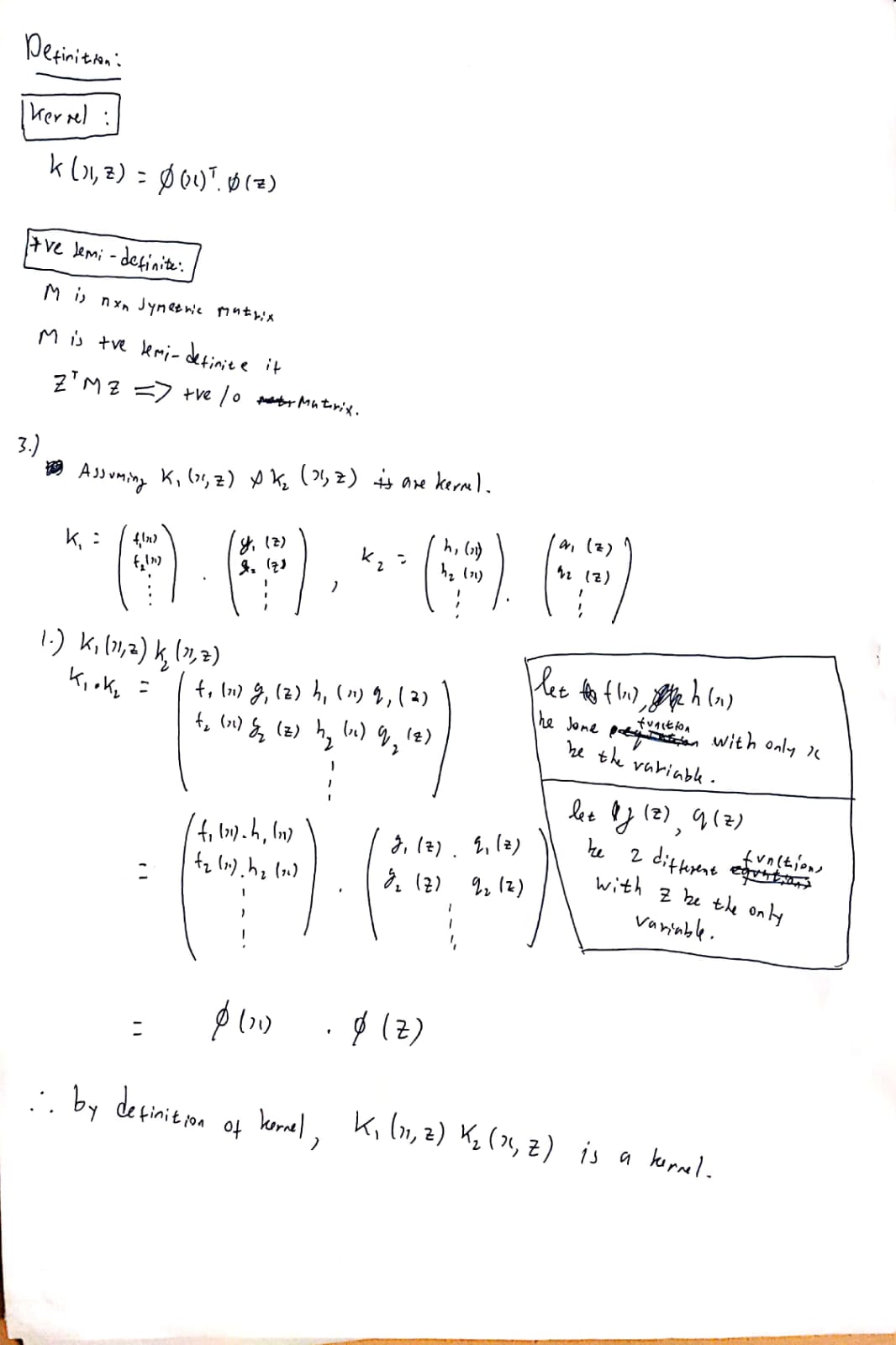
2.) Positive semi-definite is used

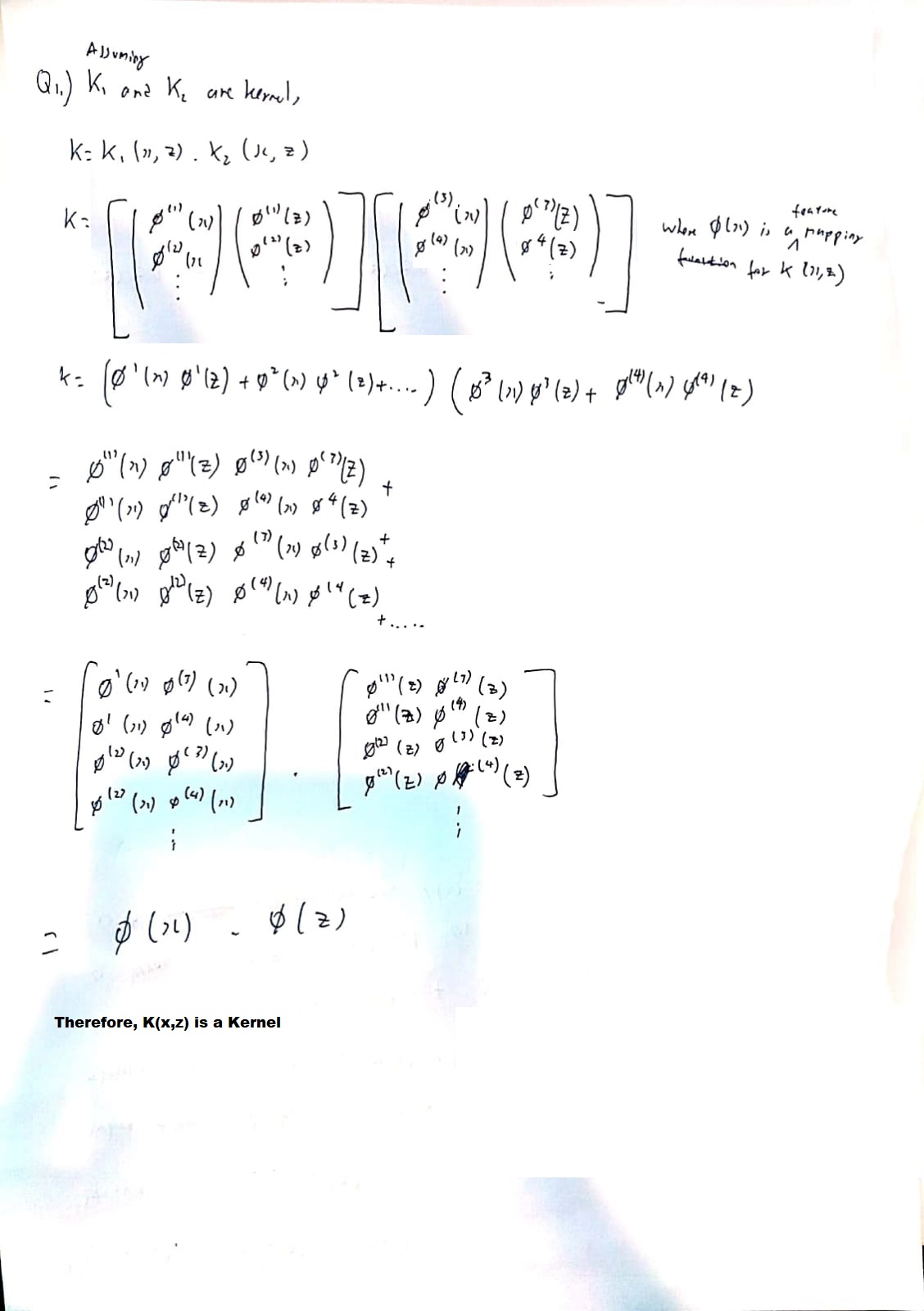
3.) Positive semi-definite is used to challenge that they are Kernel

4.) Properties of Kernel

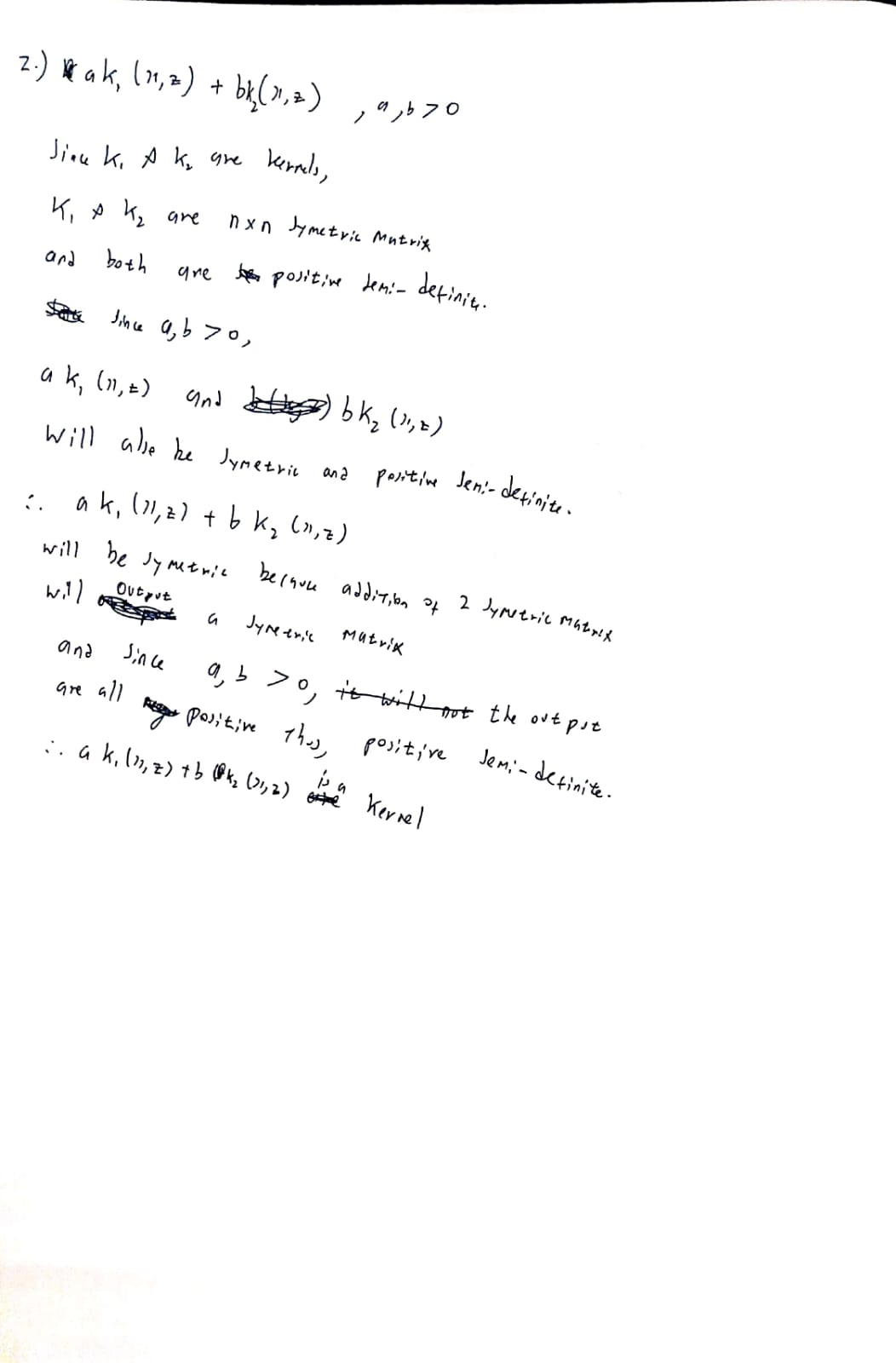
The images are also inside the folder outside of this documents.

**Qn3. Part1)**

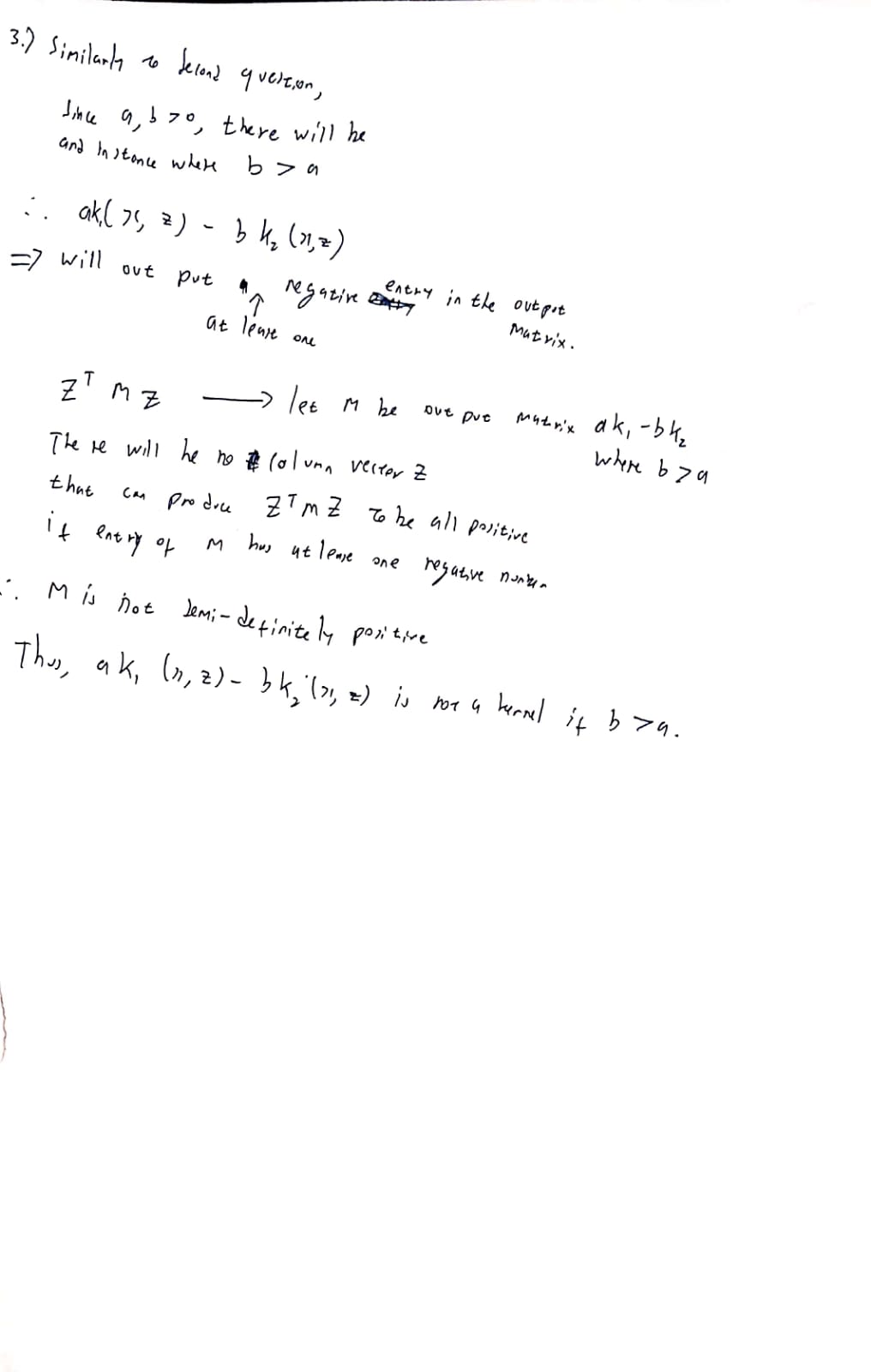
****

****

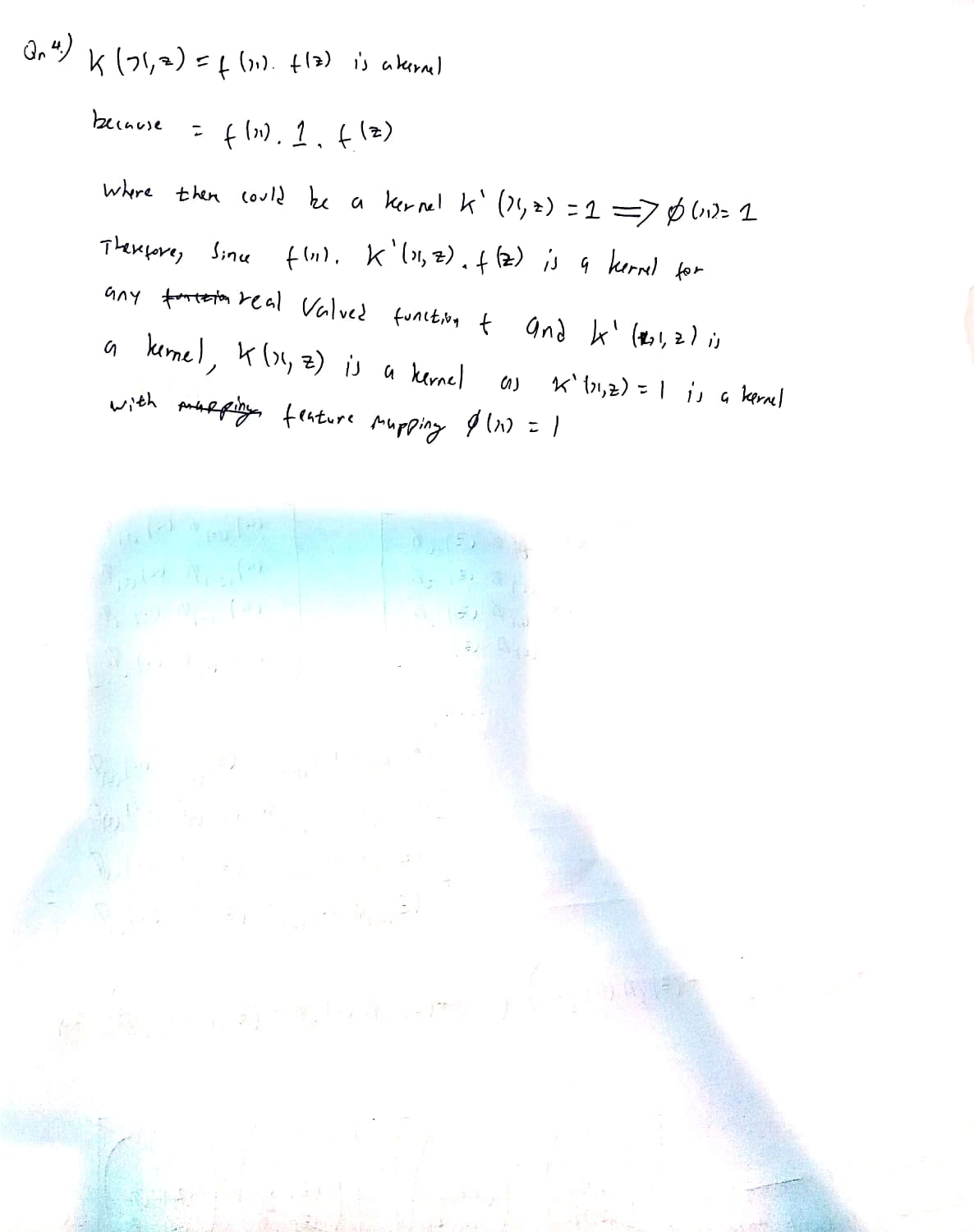
**Qn3. Part2)**

****

**Qn3. Part3)**

****

**Qn3. Part 4.)**

****

**Qn4.)**

For question 4, please refer to my folder name “Qn4” and then run the code accordingly from cell to cell. Most of the answer are already shown in the Jupyter notebook itself.