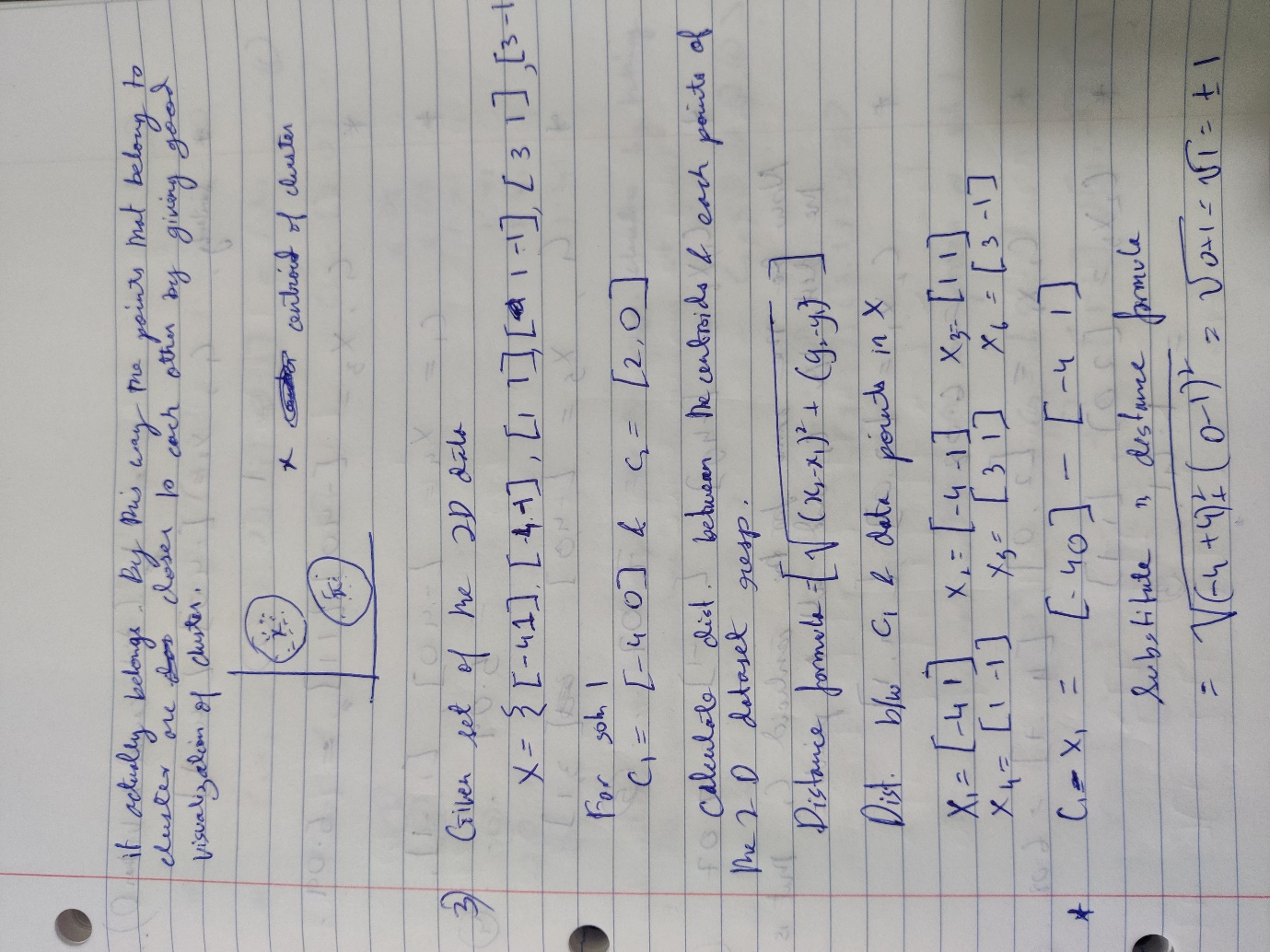
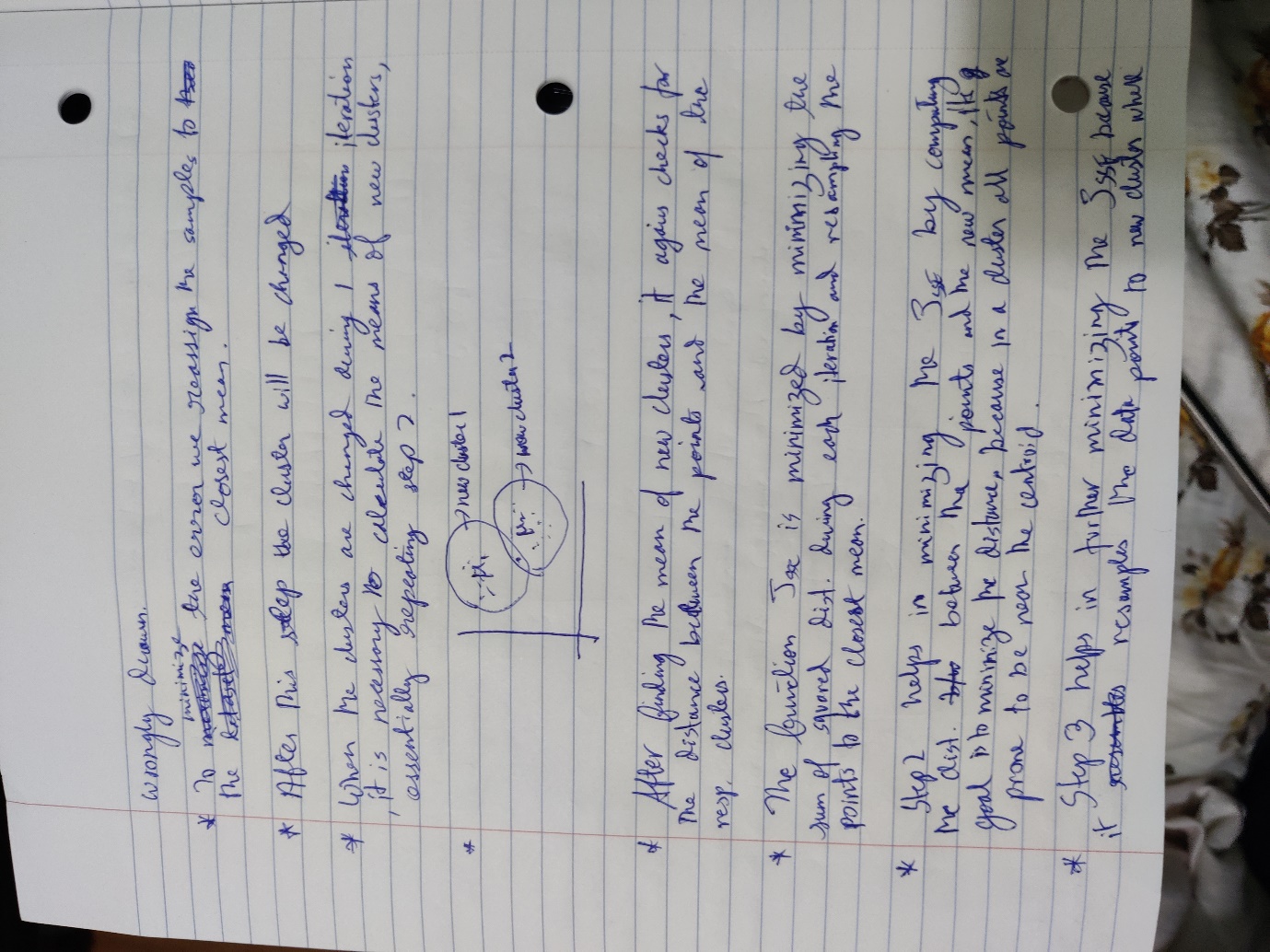
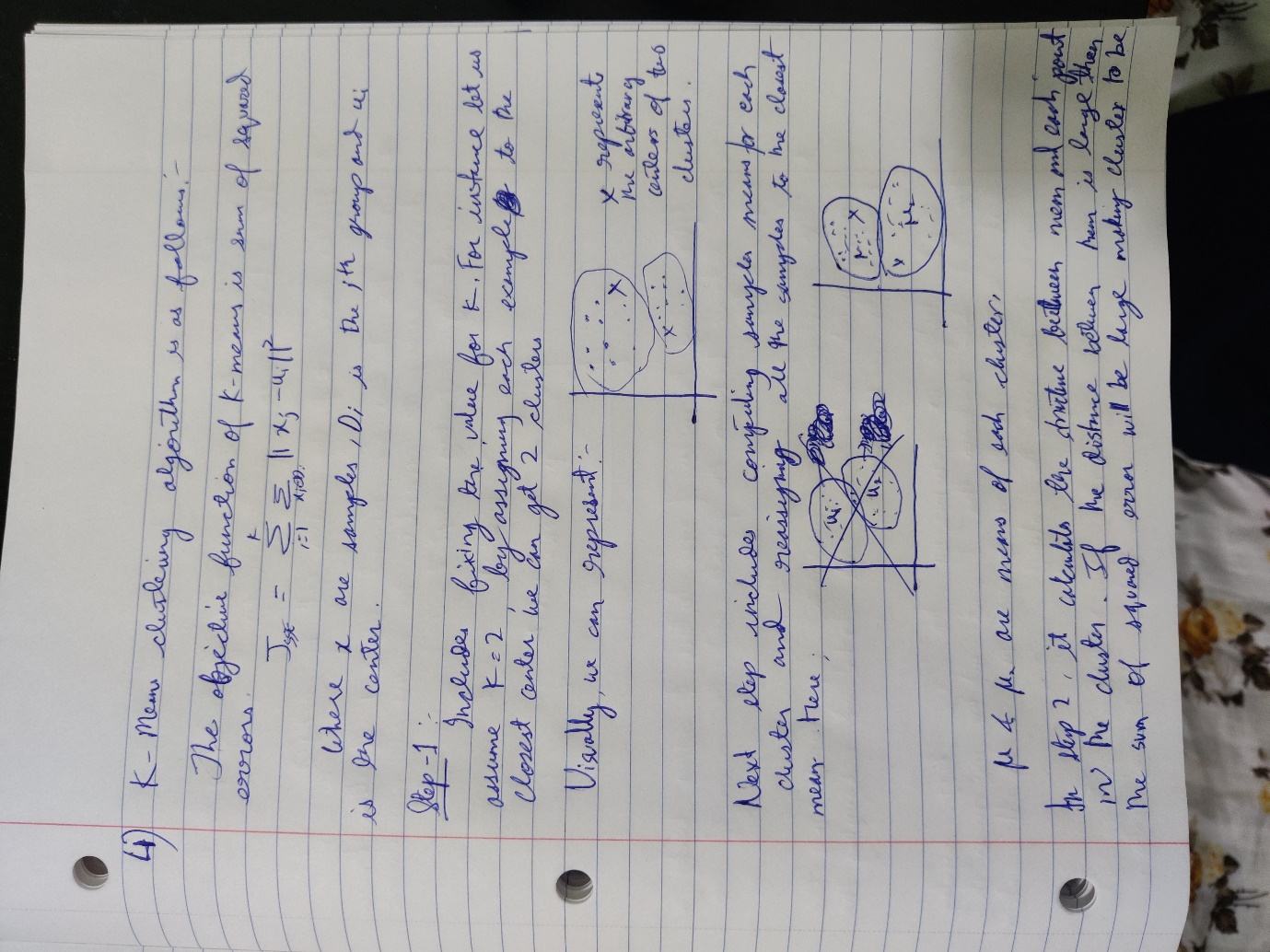
6

1. While both the Support Vector Machine(SVM) and Perceptron can be used in case of training on 2 linearly separable classes, there are certain differences and the technique of maximizing the margin using SVM is better that minimizing misclassified cases using Perceptron. The condition of stoppage of algorithm is the key difference between these two. Perceptron stops after it classifies data correctly whereas SVM stops after finding the best plane that has the maximum margin, i.e. the maximum distance between data points of both classes. Maximizing the margin distance provides some reinforcement so that future data points can be classified with more confidence.

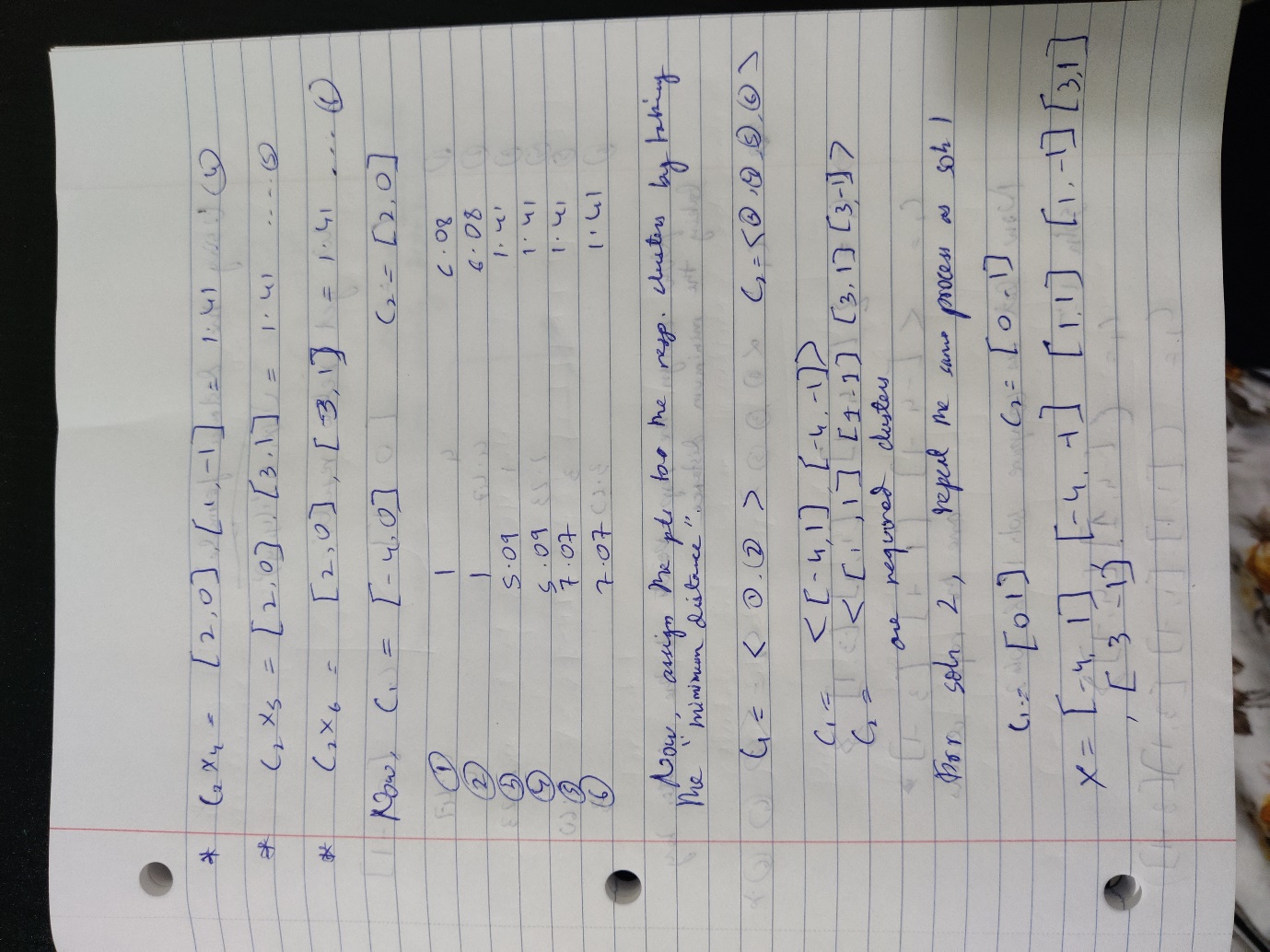
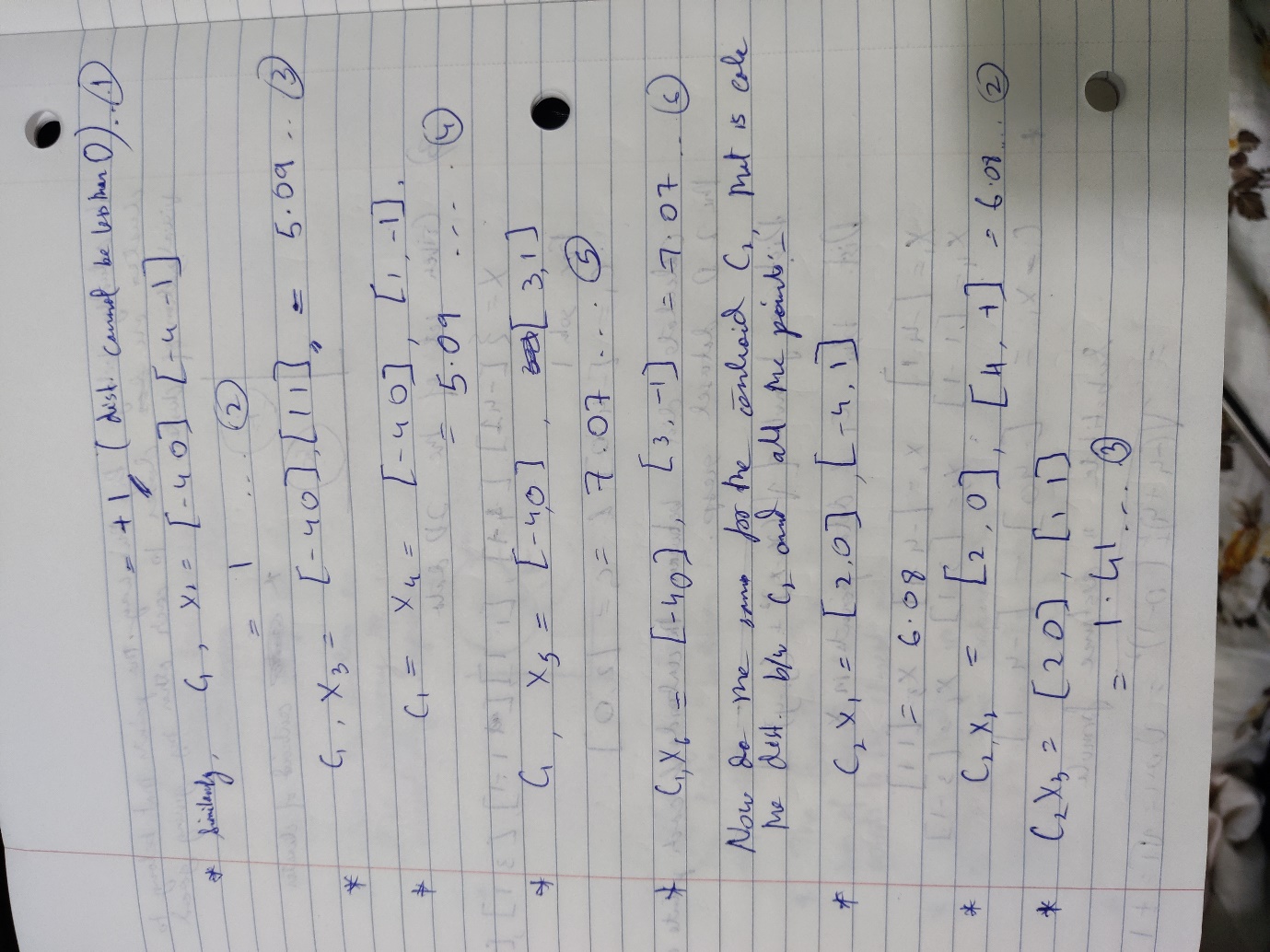
SVM keeps a classification margin on each end so that it classifies test data points that come near to the boundary properly. The margin of the SVM makes SVM more robust in getting more closer to the real boundary (target function) of the datasets.

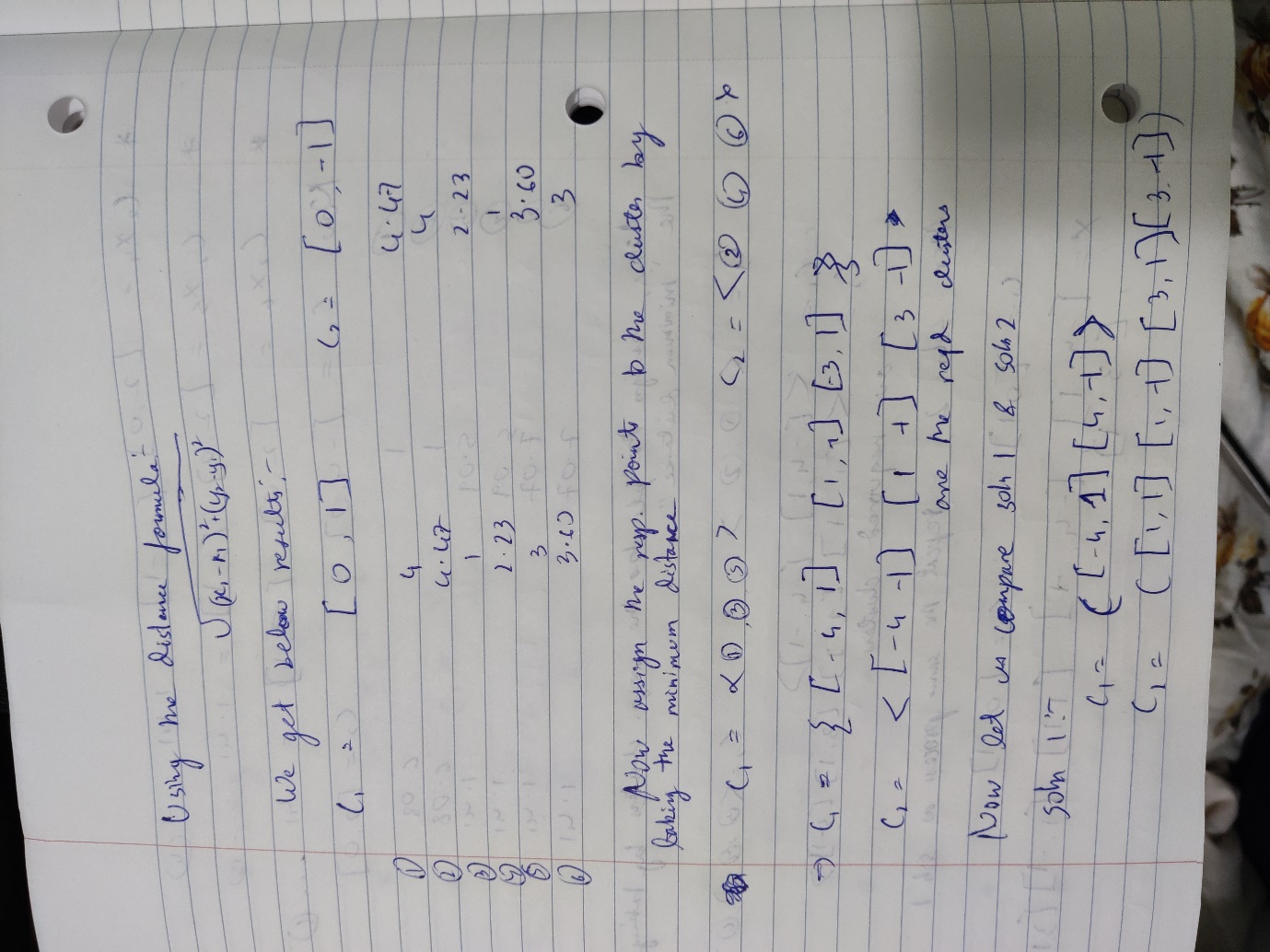
1. Based on the principles of Bayesian statistics, MLC provides a parametric approach in decision making where the model parameters need to be estimated before they are applied for classification. On the contrary, SVM is a nonparametric approach, where the theoretic background is supervised machine learning.

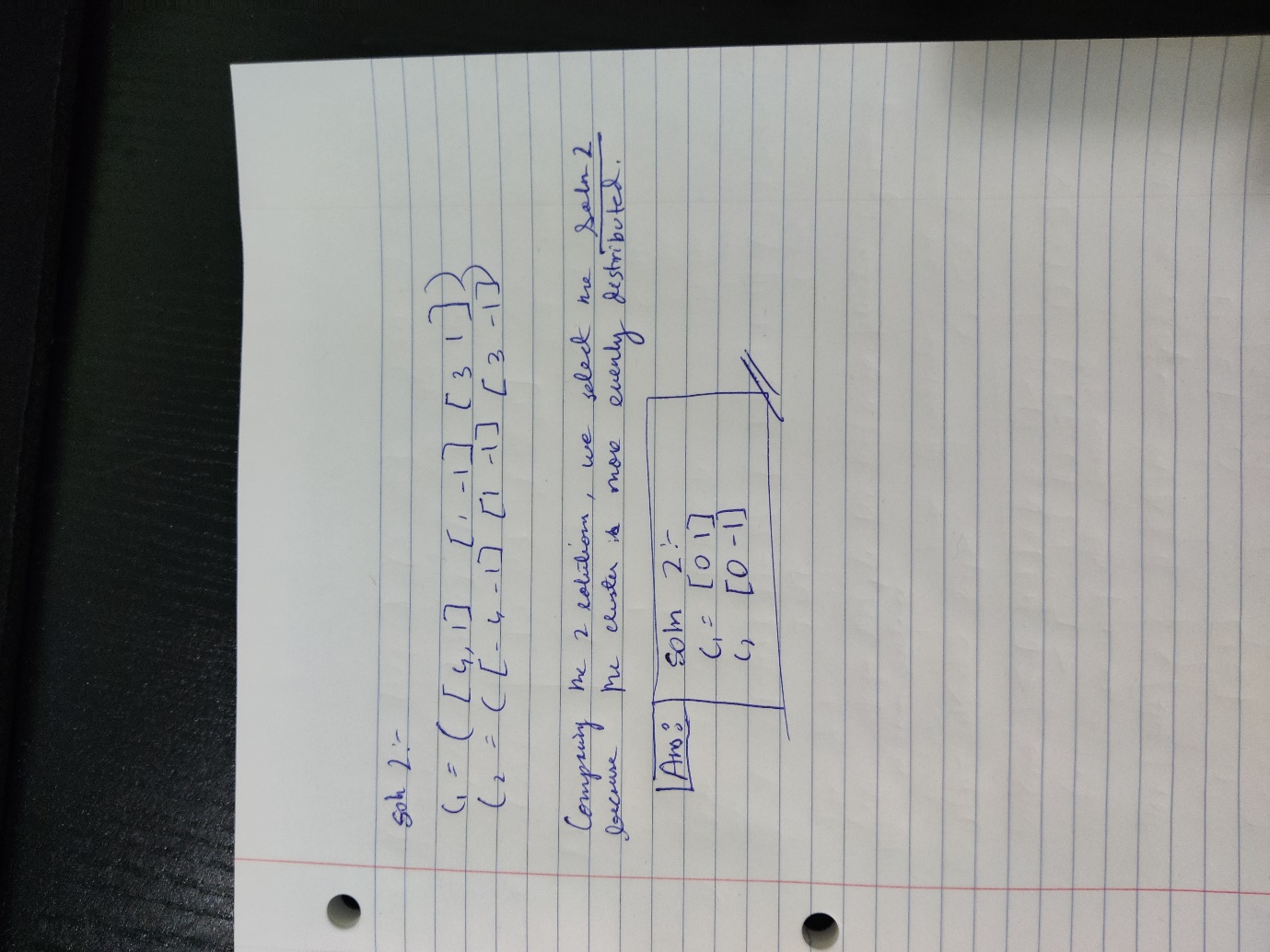
4



3) (contd: refer above)







1

1) True

2) False : random samples are chosen from the input with replacement.

3) False : Non-linear function can convert non-linear relationships to a linear

relationship at every layer. In the end, it becomes useful to draw decision boundary.

4) False : The decision regions lead to are convex too.

5) True : The testing data should be a good representation of the training data, to get accurate results.

6) True : The distance measure selection will be based on the dataset and objective. Hence,

7) False

8) True

9) False : Using slack variable will help reduce overfitting.

10) False : EM is a local optimization method. It can be at a sub-optimum, therefore, no certainty it will find global-optima.

11) True

12) True

13) True :

14) True : Initially the weights are assigned randomly. Later we can adjust the weights.

15) True

2

1. If the weak classifier does not correctly classifies a sample the weight of the sample is increased. Otherwise, if the classifier correctly classifies it, the weight of that sample decreases.
2. Weak learners are chosen on the basis of maximum error. Each sample has a weight and thus the error is higher for the sample with max. weight if misclassified. So, to pick the subsequent weaker learner, the weak learner must classify the samples with higher weight correctly else the error will be much higher and thus the weight updation affects the subsequent weak learner selection
3. In option 1, D2 is not valid as the multiplication rate/error =0 in this case.

Misclassification rate/error=( Eyi+D)/ Edi = (0.1667+0.1667+0.1667-0.25-0.25)/1

Therefore, only the weight in option 2 is valid

5

