

PATTERN RECOGNITION LAB  
CSE 4214

LAB EXPERIMENT 1

DESIGNING A MINIMUM DISTANCE TO CLASS MEAN CLASSIFIER

SUBMITTED BY

MD. MUKITUL ISLAM 14.02.04.076

SECTION: B1

26 MAY 2018



AHSANULLAH UNIVERSITY OF SCIENCE & TECHNOLOGY

# Designing a Minimum Distance to Class Mean Classifier

Md. Mukitul Islam (140204076)  
Dept. of Computer Science & Engineering  
Ahsanullah University of Science & Technology

## I. OBJECTIVE

Objective of this experiment is to design a Minimum Distance to Class Mean Classifier to classify unclassified sample vectors with the help of some given classified sample vectors.

## II. INTRODUCTION

Minimum Distance to Class Mean Classifier is used to classify unclassified sample vectors where the vectors clustered in more than one classes. For example, in a dataset of  $n$  sample vectors of dimension  $d$ , if some of the sample vectors are already clustered into classes and if some are not classified then we can use this classifier to classify the unknown vectors. In our dataset 12 samples are classified in two different classes and 4 samples are not classified. We have used this minimum distance to class mean classifier that we have implemented to classify those 4 unclassified samples.

## III. EXPERIMENTAL DESIGN AND IMPLEMENTATION

For designing a minimum distance to class mean classifier we have used some classified sample vectors and for testing purpose we have used some unclassified sample.

### A. Description of the Experimental Parameters

We have used the following two-class set of prototypes-  
 $w1 = (2,2), (3,1), (3,3), (-1,-3), (4,2), (-2,-2)$   
 $w2 = (0,0), (-2,2), (-1,-1), (-4,2), (-4,3), (2,6)$

These are the classified sample points we have used in this experiment to design this classifier.

For testing purpose we have used the following points-

$X_1 = (-1,-1)$

$X_2 = (3, 2)$

$X_3 = (-2, 1)$

$X_4 = (8, 2)$

We have classified this points either in class  $w1$  or  $w2$  based on the decision of this classifier.

### B. Class Mean & Distance Calculations

For determining class mean we have use MATLAB defined function `mean()` [1]. In this `mean()` function we pass the values of class  $w1$  and  $w2$  and it returns their class mean value respectively. In the figure red circle signifies the mean of class 1 and green circle signifies the mean of class 2. For plotting all these points we have used MATLAB defined function `plot()`. [1]

Distance calculation for a testing point from each class mean-

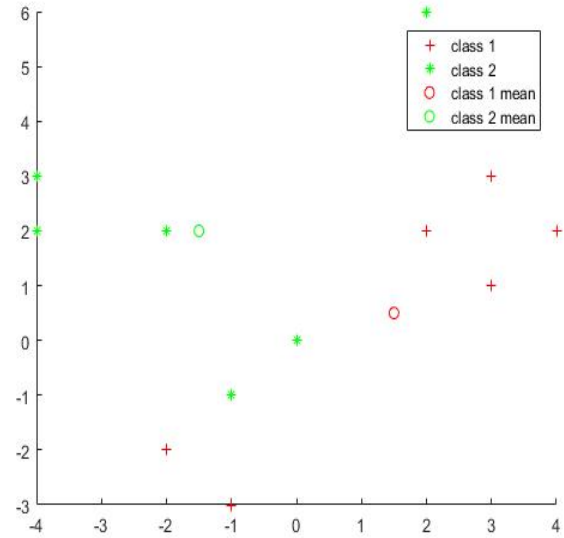


Fig. 1. Class 1 & Class 2 points with their Class Mean

Let us consider, there is  $i$  classes and  $u_1, u_2, \dots, u_i$  be the class mean of each class respectively and there is  $i$  unknown samples  $x_1, x_2, \dots, x_i$ , we have to calculate the distance from each class of these samples. For distance calculation we have used Euclidean distance calculation method. Let,  $D_i$  be the distance from  $i$ th class-

$$D_i = \sqrt{(x_1 - u_1)^2 + (x_2 - u_2)^2 + \dots + (x_i - u_i)^2}$$

$$\text{or, } D_i^2 = (x_1 - u_1)^2 + (x_2 - u_2)^2 + \dots + (x_i - u_i)^2$$

$$\text{or, } D_i^2 = (x - u)^t (x - u)$$

$$\text{or, } D_i^2 = x^t x - x^t u - u^t x + u^t u$$

$$\text{or, } D_i^2 = x^t x - 2u^t x + u^t u$$

$$\text{or, } D_i^2 = -2u^t x + u^t u$$

$$\text{or, } -\frac{1}{2}D_i^2 = u^t x - \frac{1}{2}u^t u \quad (1)$$

Here, (1) gives us the distance of any unclassified sample  $x$  from the class mean  $u$  of  $i$ th class.

### C. Decision Rule & Algorithm Implementation

In this problem there are 2 classes -  $w_1$  and  $w_2$ . We have to generate a discriminant function to make decision rule. Discriminant function for  $i$ th class-

$$g_i(x) = -\frac{1}{2}D_i^2 = u^t x - \frac{1}{2}u^t u \quad (2)$$

So, for class  $w_1$  and  $w_2$  we have got discriminant function  $g_1$  and  $g_2$  respectively. Using this discriminant function we can make the decision rule.

Here, the decision rule is-

if  $g_1(x) > g_2(x)$  then  $x \in w_1$

if  $g_1(x) < g_2(x)$  then  $x \in w_2$

Using these two decision rule this algorithm classifies any unknown sample  $x$  to class either  $w_1$  or  $w_2$ .

In this figure, sample of 12 points from class 1 and class

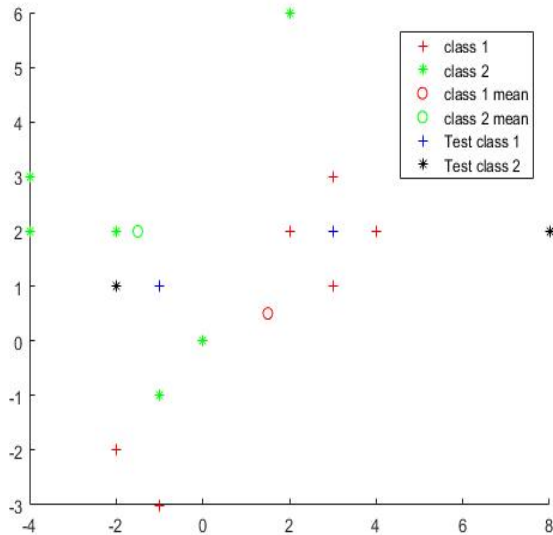


Fig. 2. Class 1 & Class 2 points with Classified Testing Points

2 are represented by '+' and '\*' marker respectively. Plot of four testing points are also in this figure. These four points are also represented with the same marker but with different color. Four testing points are classified here according to the decision rule we have implemented.

### D. Decision Boundary

The two discriminant function can be used to generate the formula for the decision boundary of the two classes  $w_1$  and  $w_2$ . For decision boundary,

$$g_1 - g_2 = 0$$

$$\text{or, } (u_1 - u_2)^t x - \frac{1}{2} \{u_1^t u_1 - u_2^t u_2\} = 0 \quad (3)$$

Here, (3) gives us the decision boundary of the two class  $w_1$  and  $w_2$ .

## IV. RESULT ANALYSIS

In our dataset there are 12 samples for training purpose these are classified samples and there are 4 samples for testing and these unclassified. From these figure we can see that among the

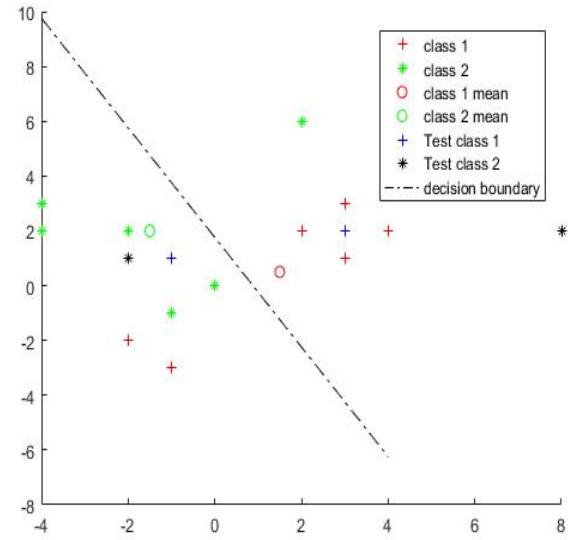


Fig. 3. Class 1 & Class 2 points with Decision Boundary

four testing points, 2 of them classified properly and for two occasions the classifier is failed and misclassified two points. So, accuracy is 50% for the dataset and testing points we have used here.

## V. CONCLUSION

This algorithm is a very simple algorithm and easy to implement and apply. As it conducts simple calculations, its calculation is apparently faster. The weakness of the algorithm is its misclassification rate is also relatively higher because the boundary between the two classes is linear.

## REFERENCES

- [1] MATLAB Documentations, <https://www.mathworks.com/>