<u>CMSC 691: Intro. to Data Science</u>
<u>HW – 3</u>

<u>AKSHAY PESHAVE</u>
(peshave1@umbc.edu)

QUESTION 1

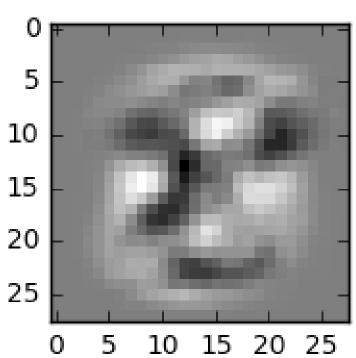
Logistic Regression Parameters:

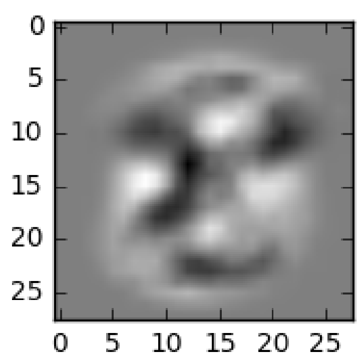
Learning Rate Decay: Exponential function
Learning Iterations : 50
Training Set: Entire dataset (10,000 observations, 28x28 features)

Learning Rate	**Decay Step Value**	**Dataset Accuracy**
0.0001	No Decay	91.98%
0.1	5	91.88%
0.1	**No Decay**	**93.67%**
0.5	5	91.92%
0.5	10	93.00%

Below is the weight vector (for learning rate =0.1 with no decay) mapped as a 28x28 pixel image. The image on the left is drawn with nearest value color interpolation. This makes it appear pixelated but gives us a better idea of what the regressor learned. We observe that the model predicts the class based mainly on:

1. Left side concave curve of the number ‘8’ (in dark black). Priority is given to the lower part of this concavity since the upper part applies to the number ‘5’ as well.
2. The voids inside the ‘8’ and on the sides of the point of intersection of the left and right curves (in bright white). All these voids occurring together with the above feature set is highly indicative of the number being an ‘8’.
3. The top right curvature is also given significance. But it is relatively less (in dark gray) since that part applies to the number ‘2’ as well





The image on the right is the same pixel map rendered using bilinear color interpolation to get a smoother image. This shows the above weight effects (resulting in the shape of the number ‘8’ with appropriate voids) more prominently.
The area surrounding the number is of around 50% gray shade. This is because the data has no variation for those pixels since the numbers are drawn at the center of the image.

QUESTION 2

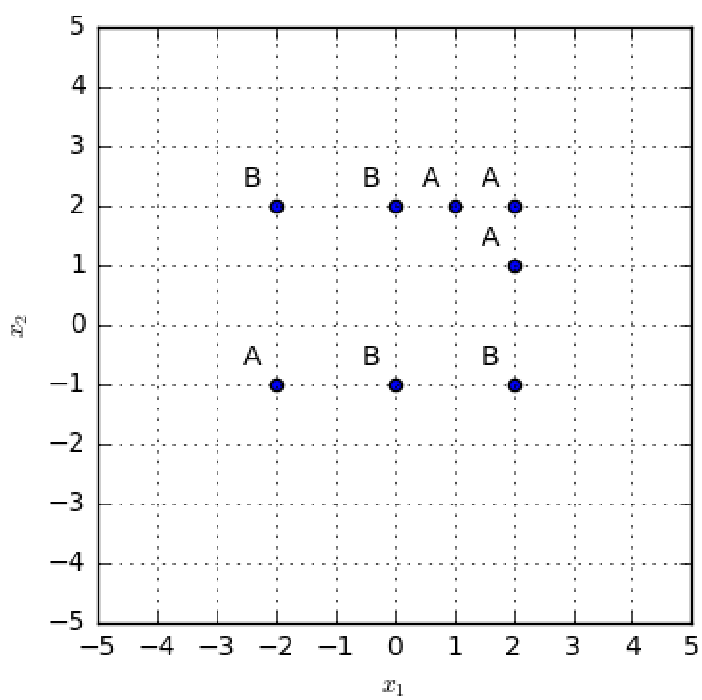
INPUT SPACE

VEC #	x_1	x_2	class
1	2	2	A
2	-2	-1	A
3	1	2	A
4	2	1	A
5	-1	2	B
6	0	2	B
7	0	-1	B
8	2	-1	B

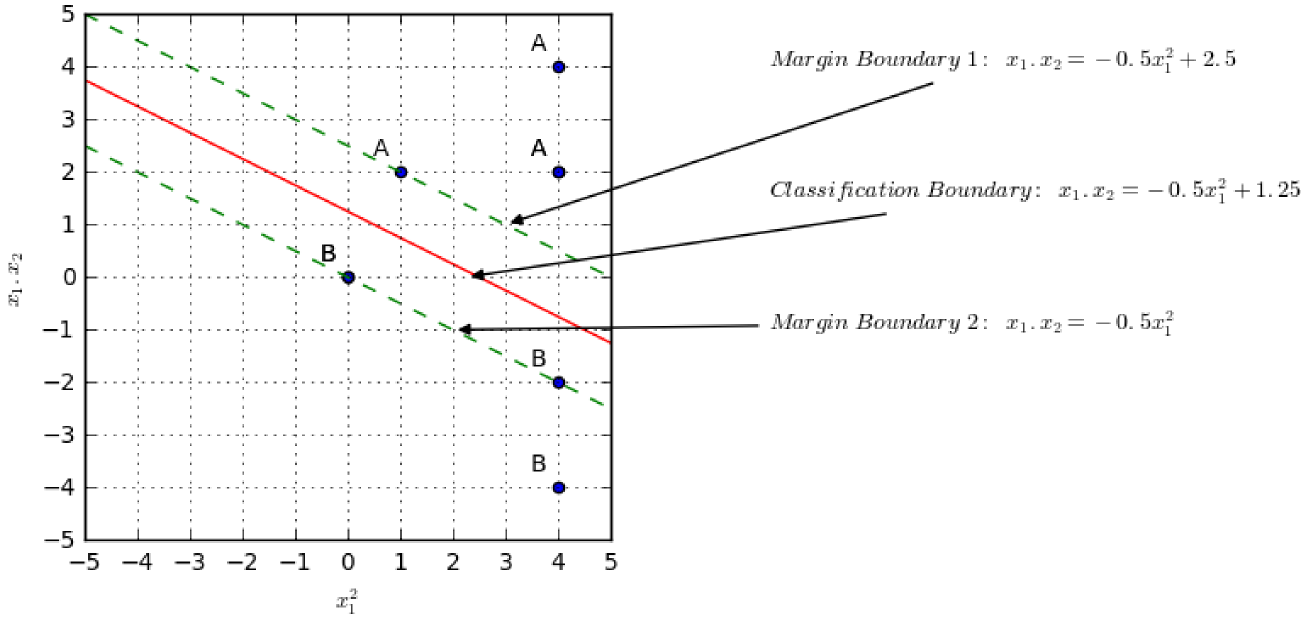
FEATURE SPACE

VEC #	x_1^2	$x_1.x_2$	class
1	4	4	A
2	4	2	A
3	1	2	A
4	4	2	A
5	4	-4	B
6	0	0	B
7	0	0	B
8	4	-2	B

(a)



(b) and (c)



The support vectors in the feature space are:

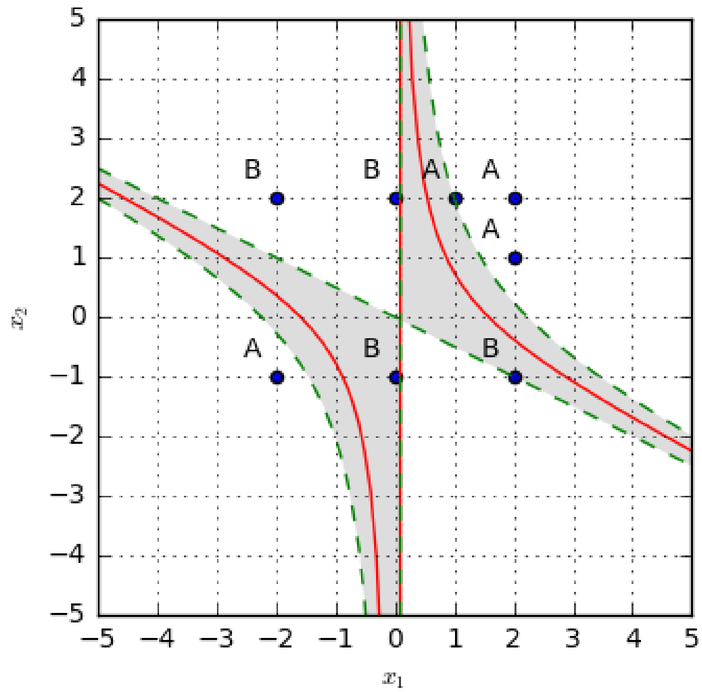
VEC #	X1^2	X1.X2	class
3	1	2	A
6	0	0	B
7	0	0	B
8	4	-2	B

(d)

The support vectors in the input space are:

VEC #	X1	X2	class
3	1	2	A
6	0	2	B
7	0	-1	B
8	2	-1	B

(e) and (f)



(g)

The equation of the linear separator in the feature space is :

$$x_1.x_2 = -0.5x_1^2 + 1.25 \quad \text{or} \quad x_1.x_2 + 0.5x_1^2 = 1.25$$

All points which lie in the region of class ‘A’ are the right of the linear separator in the feature space. These points, in the feature space, will necessarily lie on a line from the set of lines represented by the inequality : $x_1.x_2 + 0.5x_1^2 > 1.25$

Thus, if $x_1.x_2 + 0.5x_1^2 > 1.25$ then the vector belongs to class ‘A’ otherwise to class ‘B’.

A simpler inequality may be obtained if we use the margin boundary of class ‘B’. This margin boundary is represented in the feature space by the line : $x_1.x_2 = -0.5x_1^2$ or $x_2 = -0.5x_1$

Using the same logic as in the earlier case we can infer that, if $x_2 + 0.5x_1 > 0$ then the vector belongs to class ‘A’ otherwise to class ‘B’.