Theodore Kim
CS-UY 4563 – Introduction to Machine Learning
HW #5b

1. Using soft-margin SVM:

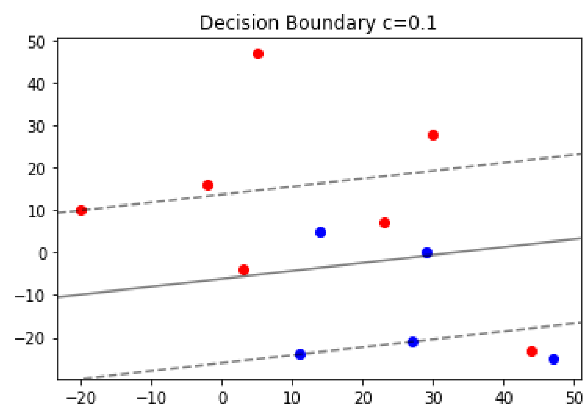
(a) If the cost (ξ_1) of a point x_i is 0, the point is located on the outside of the margin and is correctly classified.

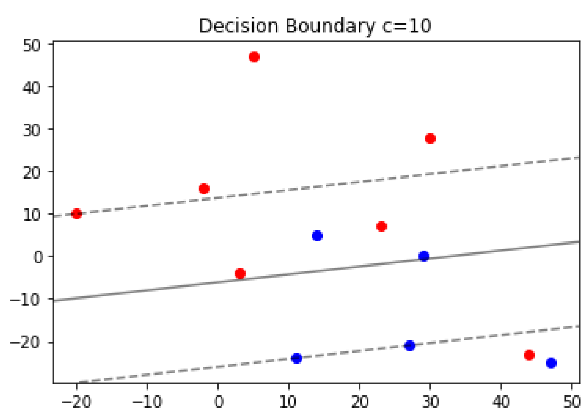
(b) If the cost (ξ_1) of a point x_i is greater than 0 and less than 1, the point is located within the margin but on the "correct" side of classifier line, therefore the classification was "correct" but there is a chance that, the relaxation of the margin misclassified the point.

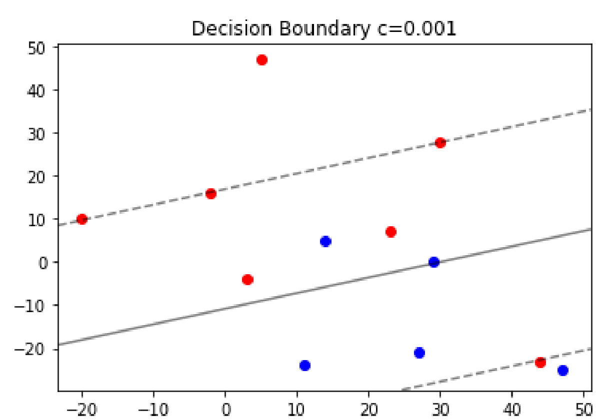
(c) If the cost (ξ_1) of a point x_i is greater than 1, the point is located on the wrong side of the classifier line and the margin and is therefore incorrectly classified.

2. Plot the decision boundary when C=0.1 and C=10

x_1	x_2	y
5	47	-1
14	5	1
47	-25	1
3	-4	-1
-2	16	-1
30	28	-1
27	-21	1
11	-24	1
29	0	1
23	7	-1
-20	10	-1
44	-23	-1





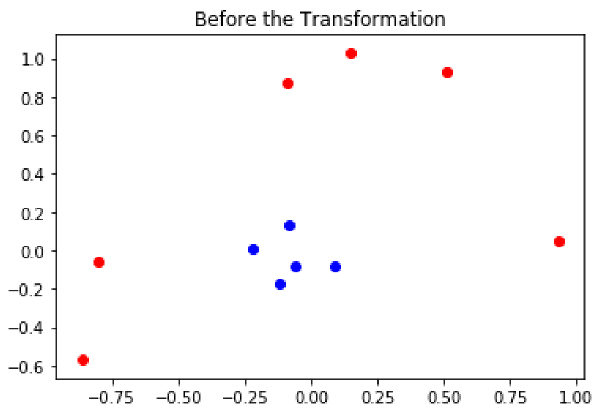


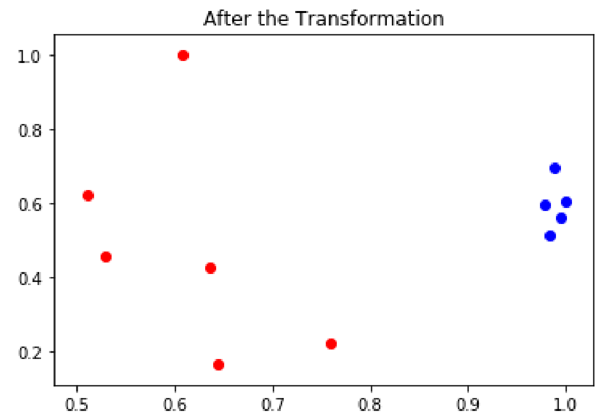
3. Transform the below dataset in order to make it linearly separable.

a) For the below data, I used an RBF kernel transformation using the landmark set:

$l = \{(-0.06, 0.08), (0.93, 0.05)\}$ These points were chosen for their relatively central position to the two class clusters.

x_1	x_2	z_1	z_2	y
0.15	1.03	0.5283	0.4564	-1
-0.08	0.13	0.9780	0.5985	1
-0.8	-0.06	0.7603	0.2226	-1
-0.22	0.01	0.9833	0.5158	1
0.51	0.93	0.5104	0.6216	-1
-0.09	0.87	0.6365	0.4247	-1
0.09	-0.08	0.9888	0.6968	1
-0.12	-0.17	0.9942	0.5625	1
0.93	0.05	0.6074	1.0000	-1
-0.06	-0.08	1.0000	0.6074	1
-0.86	-0.57	0.6440	0.1663	-1

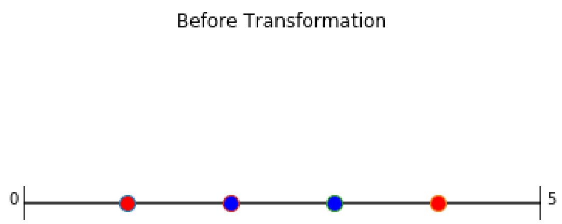


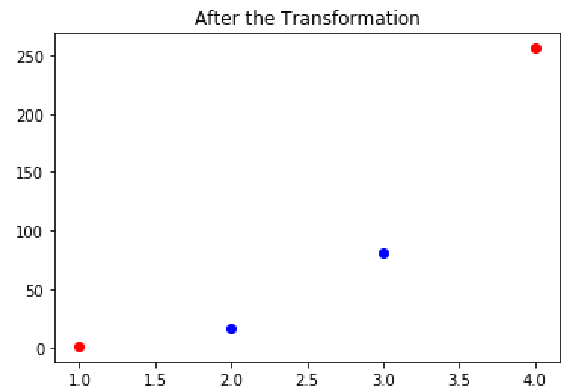


b) For the below data

x_1	y
1	-1
2	1
3	1
4	-1

I made it linearly separable using a polynomial kernel transformation with degree = 4. I transformed the data into two features: $\phi(x) = [x, x^2]$





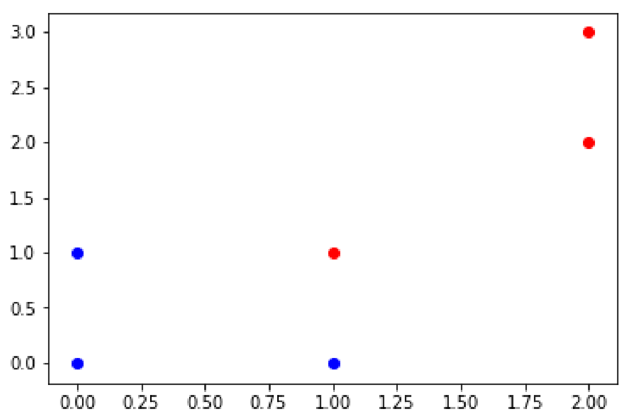
4. If in question (3A), if the point (-0.08, 0.13) is removed, does the margin stay the same? Which points, if removed, will change the margin?

(-0.18, 0.13) would not change the margin as its transformed feature (0.9780, 0.5985) does not lie on the support vector. On the other hand, the point (-0.22, 0.01) would change the margin as its transformed feature, (0.9833, 0.5158), does lie on the support vector. Finally, the points (-0.8, -0.06) and (0.93, 0.05) lie on opposite support vector, and removing them would also alter the margin.

5. Given the data:

x_1	x_2	y
1	1	-1
2	3	-1
2	2	-1
0	0	1
1	0	1
0	1	1

a) Plot the points and determine if they are separable.



They do seem linearly separable.

b) Write the constraints for the constrained optimization problem:

$\min w_1^2 + w_2^2$

subject to $-1(w_0 + w_1 + w_2) \geq 1$

$-1(w_0 + 2w_1 + 3w_2) \geq 1$

$-1(w_0 + 2w_1 + 2w_2) \geq 1$

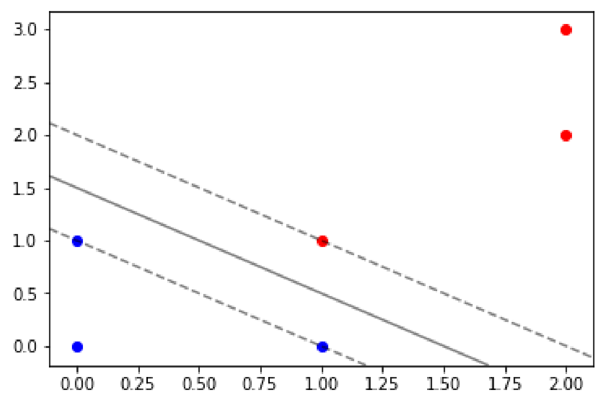
$1(w_0) \geq 1$

$1(w_0 + w_1) \geq 1$

$1(w_0 + w_2) \geq 1$

c) Find the hyperplane that separates the points and has the maximal margin

$$w = \begin{bmatrix} -1.99921875 \\ -1.99921875 \end{bmatrix} \; w_0 = 2.99895833$$



d) Identify the support vectors

The support vectors are:

(0, 1) and (1,0)
(1,1)

<u>6. Demonstrate if $\phi(x) = \phi([x_{i1}, x_{i2}]) = (1, \sqrt{2}x_{i1}, \sqrt{2}x_{i2}, x_i^2, x_2^2, \sqrt{2}x_{i1}x_{i2})$, then $\phi(x)^T\phi(x') = (1 + x^T x)^2$</u>

$$\phi(x)^T = \phi([x_{i1}, x_{i2}])^T = \begin{bmatrix} 1 \\ \sqrt{2}x_{i1} \\ \sqrt{2}x_{i2} \\ x_1^2 \\ x_2^2 \\ \sqrt{2}x_{i1}x_{i2} \end{bmatrix}$$

$$\phi(x') = \phi([x_{i1}, x_{i2}]) = \begin{bmatrix} 1 & \sqrt{2}x_{1i} & \sqrt{2}x_{2i} & x_i^2 & x_{i+1}^2 & \sqrt{2}x_{1i}x_{2i} \end{bmatrix}$$

$$\phi(x)^T \cdot \phi(x') = \begin{bmatrix} 1 \\ \sqrt{2}x_{i1} \\ \sqrt{2}x_{i2} \\ x_1^2 \\ x_2^2 \\ \sqrt{2}x_{i1}x_{i2} \end{bmatrix} \cdot \begin{bmatrix} 1 & \sqrt{2}x_{1i} & \sqrt{2}x_{2i} & x_i^2 & x_{i+1}^2 & \sqrt{2}x_{1i}x_{2i} \end{bmatrix}$$

$$\begin{aligned}\phi(x)^T \cdot \phi(x') = \Big(1 &+ \left(\sqrt{2}x_{i1} \cdot \sqrt{2}x_{1i}\right) + \left(\sqrt{2}x_{i2} \cdot \sqrt{2}x_{2i}\right) + \left(x_1^2 \cdot x_i^2\right) + \left(x_2^2 \cdot x_{i+1}^2\right) \\ &+ \left(\sqrt{2}x_{i1}x_{i2} \cdot \sqrt{2}x_{1i}x_{2i}\right)\Big)\end{aligned}$$

$$\phi(x)^T \cdot \phi(x') = (1 + \ x^T x)^2$$