Introduction to Machine Learning
Homework 5b: SVM

Prof. Linda Sellie

1. In class we discussed the soft-margin SVM:

$$\min \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i$$

$$\texttt{subject to } y_i(w^T x_i + w_0) \geq 1 - \xi_i, \texttt{ and } \xi_i \geq 0$$

   (a) For a point x_i if $\xi_i = 0$, what do we know about the where x_i is wrt the margin. Is x_i correctly classified by the hyperplane?

   (b) For a point x_i if $0 < \xi_i \leq 1$, what do we know about the where x_i is wrt the margin. Is x_i correctly classified by the hyperplane?

   (c) For a point x_i if $\xi_i > 1$, what do we know about the where x_i is wrt the margin. Is x_i correctly classified by the hyperplane?

2. Given the following points, plot the decision boundary when $C = 0.1$ and $C = 10$. You can use the code from https://jakevdp.github.io/PythonDataScienceHandbook/05.07-support-vector-machines.html to help you draw the decision boundary.

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. For the following non linearly separable points, find a transformation to make them linearly separable. Using matplotlib to plot the points before and after your transformation.

   (a)

x_1	x_2	y
0.15	1.03	-1
-0.08	0.13	1
-0.80	-0.06	-1
-0.22	0.01	1
0.51	0.93	-1
-0.09	0.87	-1
0.09	-0.08	1
-0.12	-0.17	1
0.93	0.05	-1
-0.06	-0.08	1
-0.86	-0.57	-1

   (b)

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

4. For the points in question **??**, if you remove $(-0.08, 0.13)$, does the margin stay the same? Describe which points will not change the margin and which points might change the margin.

5. Given the following points:

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

   - Plot the 7 training points. Are they separable?
   - Write the constraints for the constrained optimization problem to maximize the margin that separates the points
   - Find the hyperplane which has the maximal margin and separates the points
   - Identify the support vectors

6. Show that if $\phi(\mathbf{x}) = \phi([x_{i1}, x_{i2}]) = (1, \sqrt{2}x_{i1}, \sqrt{2}x_{i2}, x_1^2, x_2^2, \sqrt{2}x_{i1}x_{i2})$ then $\phi(\mathbf{x})^T\phi(\mathbf{x}') = (1 + \mathbf{x}^T\mathbf{x})^2$