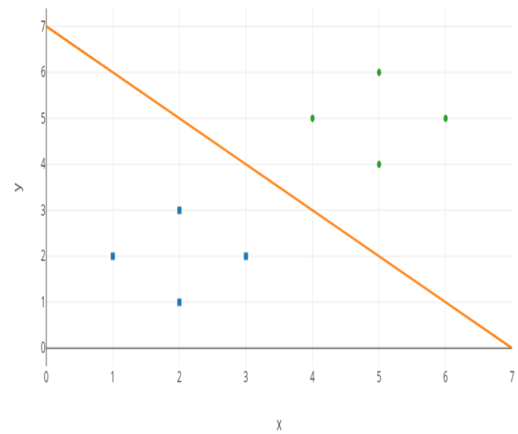
CSE 151: Homework 7 Solutions

1 Linear classification

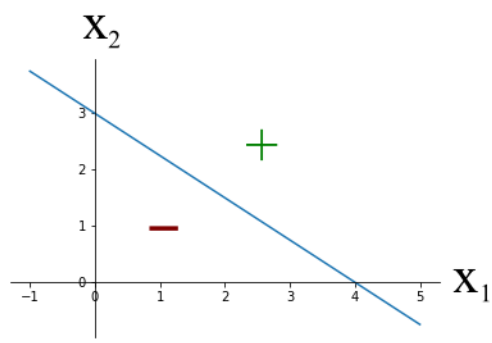
1. (a) Decision boundary is shown in figure.



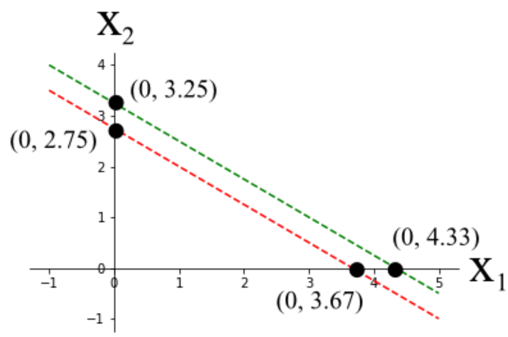
   (b) The margin is $\sqrt{2}$.

   (c) w lies in direction of $\begin{pmatrix}1\\1\end{pmatrix}$ and has length $1/\sqrt{2}$ (since the margin is $\sqrt{2}$; therefore, $w = \begin{pmatrix}1/2\\1/2\end{pmatrix}$. We know that point $x_0 = (4,3)$ lies on decision boundary; solving $w \cdot x_0 + b = 0$ yields $b = -7/2$.

2. (a) The decision boundary plot should look something like the plot below.



   (b) The left- and right-hand boundary plot should look something the plot below.



(c) The margin of this classifier is

$$\gamma = \frac{1}{\|w\|} = \frac{1}{\sqrt{3^2 + 4^2}} = \frac{1}{5}.$$

(d) The point $x = (2, 2)$ satisfies

$$w \cdot x + b = 6 + 8 - 12 = 2 > 0$$

Thus this point would be classified as $+1$.

(e) The support vectors are the points x such that $x \cdot w + b = \pm 1$. We are told that there are two distinct support vectors $x^{(1)}, x^{(2)} \in \mathbb{R}^2$ and they both satisfy $x_1^{(1)} = 1 = x_1^{(2)}$. Then it must be the case that $x^{(1)} \cdot w + b = +1$ and $x^{(2)} \cdot w + b = +1$. Solving for $x_2^{(1)}$ gives us

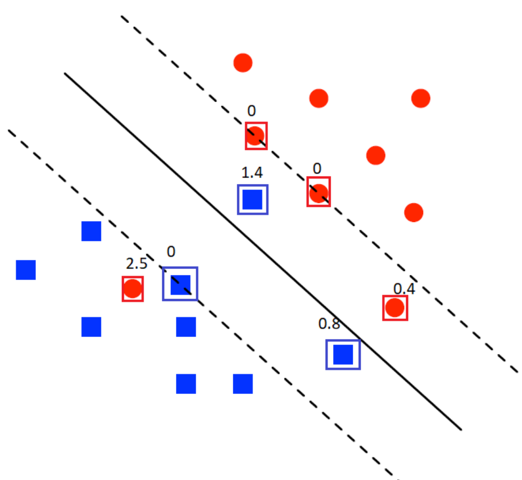
$$1 = x^{(1)} \cdot w + b = 3 + 4x_2^{(1)} - 12 = 4x_2^{(1)} - 9.$$

Thus $x^{(1)} = (1, 2.5)$. Solving for $x_2^{(2)}$ gives us

$$-1 = x^{(2)} \cdot w + b = 3 + 4x_2^{(2)} - 12 = 4x_2^{(2)} - 9.$$

Thus $x^{(2)} = (1, 2)$.

3. (a) Support vectors and their respective slack variables are marked in figure.



(b) The margin decreases if the factor C is increased.

4. To classify a point x, we evaluate the three linear functions and pick the one with the highest value.

The region where class 1 beats class 2 is:

$$w_1 \cdot x + b_1 > w_2 \cdot x + b_2 \iff (w_1 - w_2) \cdot x + (b_1 - b_2) > 0 \iff x_2 > 1$$

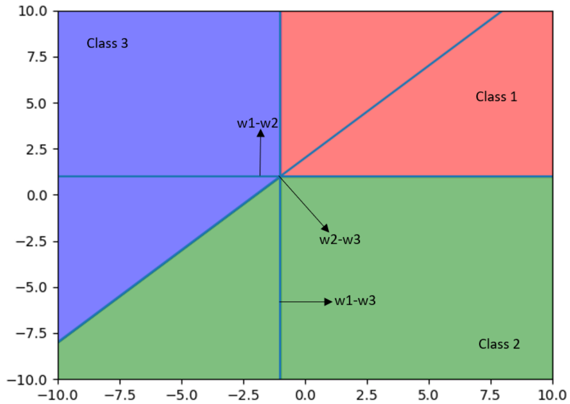
The region where class 1 beats class 3 is:

$$w_1 \cdot x + b_1 > w_3 \cdot x + b_3 \iff (w_1 - w_3) \cdot x + (b_1 - b_3) > 0 \iff x_1 > -1$$

The region where class 2 beats class 3 is:

$$w_2 \cdot x + b_2 > w_3 \cdot x + b_3 \iff (w_2 - w_3) \cdot x + (b_2 - b_3) > 0 \iff x_1 - x_2 > -2$$

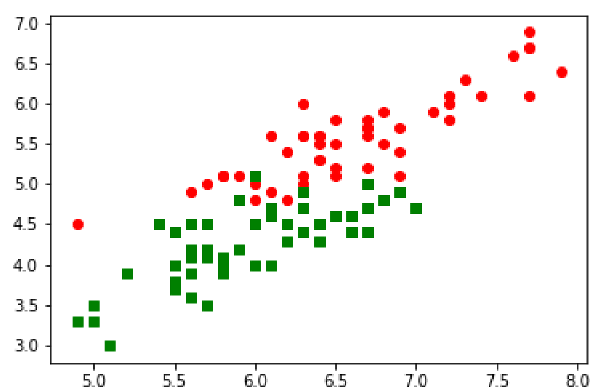
So class 1 is predicted in the intersection of the first two regions, etc. This is summarized in the figure below.



2 Programming problems

1. *Support vector machine.*

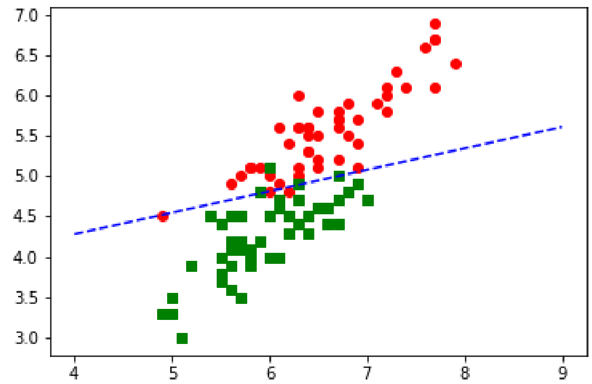
(a) The data is not linearly separable. We can see this by inspecting the scatter plot.



(b) The table you produce should look something like the following.

C value	1.5	3.0	4.5	6.0	7.5	9.0	10.5	12.0	13.5	15.0
Training error	0.07	0.05	0.06	0.05	0.05	0.05	0.05	0.04	0.05	0.07
# of support vectors	27	22	21	19	19	19	18	17	16	16

(c) The boundary plot should look something like the following.



2. *Multiclass Perceptron.*