SVM Homework

Questions?

Exercises

1. You are given a training dataset, as shown in Fig 1. Note that the training data comes from sensors which can be error-prone, so you should avoid trusting any specific point too much. For this problem, assume that we are training an SVM with a quadratic kernel.

   (a) Where would the decision boundary be for very large values of C (i.e., $C \to \infty$)? Draw on figure and justify your answer.

   (b) For $C \approx 0$, indicate in the figure where you would expect the decision boundary to be? Justify your answer.

   (c) Which of the two cases above would you expect to work better in the classification task? Why?

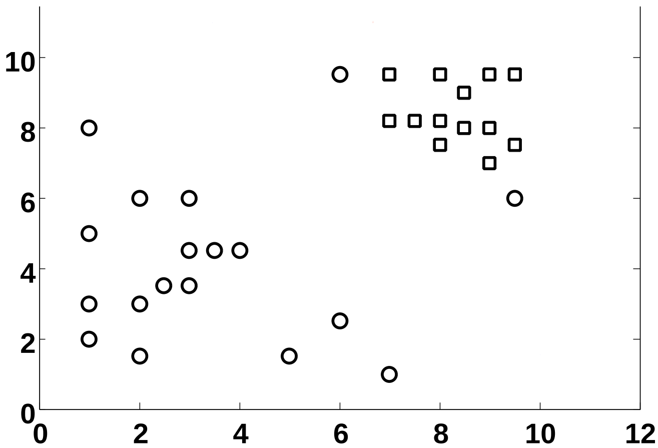

Figure 1: Training Data

2. Recall the formulation of Hard (Linear) SVM:

$$\min_{w} \frac{1}{2} w^T w$$

s.t $\forall i, y_i w^T x_i \geq 1$

(a) Complete the formulation of soft SVM:

$$\min_{w,\xi_i} \frac{1}{2} w^T w + C \sum \xi_i$$

s.t________________________________

____

(b) Complete: If C = ____________, soft SVM will behave exactly as hard SVM.

(c) In order to reduce over-fitting, one should *decrease*/*increase* the value of C. (circle the correct answer). Briefly justify your answer?

(d) We are given the dataset in Figure 1 below, where the positive examples are represented as black circles and negative points as white squares. (The same data is also provided in Table 1 for your convenience).

Recall that the equation of the separating hyperplane is $\hat{y} = w^T x - \theta$.

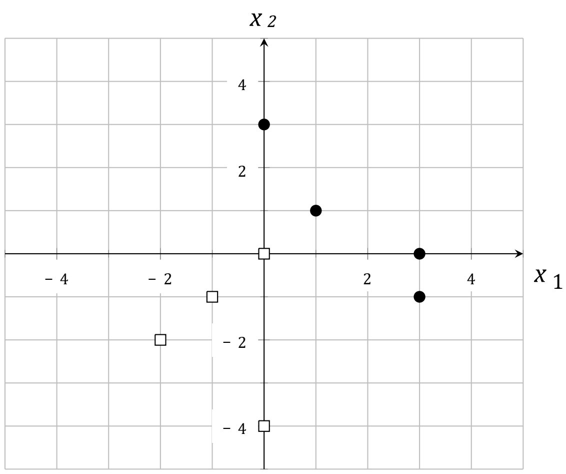

Figure 1: Linear SVM

	Attributes		
	x_1	x_2	y
1	0	0	(-)
2	0	-4	(-)
3	-1	-1	(-)
4	-2	-2	(-)
5	3	0	(+)
6	0	3	(+)
7	1	1	(+)
8	3	-1	(+)

Table 1: The data set S.

(i) **Draw** the hard SVM decision boundary for the dataset in Figure 1. **Write down** the parameter for the learned linear decision function.

$$W = (w_1, w_2) = \_\_\_\_\_\_\_\_\_\_\_\_. \ \theta =$$

(ii) **Circle** the support vectors in Figure 1.