HW3

1. We consider here linear and non-linear support vector machines (SVM) of the form:

$$\min \quad \frac{1}{2}w_1^2 \quad \text{st} \quad y_i(w_1 x_i + w_0) - 1 \geq 0, i = 1, ..., n, \text{or} \tag{1}$$

$$\min \quad \frac{1}{2}\mathbf{w}^T\mathbf{w} \quad \text{st} \quad y_i(\mathbf{w}^T \Phi_i + w_0) - 1 \geq 0, i = 1, ..., n \tag{2}$$

where ϕ_i is a feature vector constructed from the corresponding real valued input x_i. We wish to compare the simple linear SVM classifier $(w_1 x + w_0)$ and the non-linear classifier $\mathbf{w}^T\Phi + w_0$, where $\phi = [x, x^2]^T$.

1) Provide three input points x_1, x_2, and x_3 and their associated ± 1 labels such that they cannot be separated with the simple linear classifier, but are separable by the non-linear classifer with $\phi = [x, x^2]^T$. You may find Figure 1 helpful in answering this question.

2) In Figure 1, mark your three points x_1, x_2, and x_3 as points in the feature space with their associated labels. Draw the decision boundary of the SVM classifier with $\phi = [x, x^2]^T$ that separates the points.

3) Consider two labeled points $(x = 1, y = 1)$ and $(x = 3, y = -1)$. The margin we attain using feature vectors $\phi = [x, x^2]^T$ is ( ) greater ( ) equal ( ) smaller than the margin resulting from using the input x directly. (Provide a brief explanation.)

2. Support vector machines learn a decision boundary leading to the largest margin from both classes.You are training SVM on a tiny dataset with 4 points show in Figure 2. This dataset consists of two examples with class label -1(denoted with plus), and two examples with class label $+1$(denoted with triangles).

1) Find the weight vector w and bias b.What's the equation corresponding to the decision boundary?

2) Circle the support vectors and draw the decision boundary.

3. Derive the feature mapping function $\phi(x)$ for Gaussian kernel.

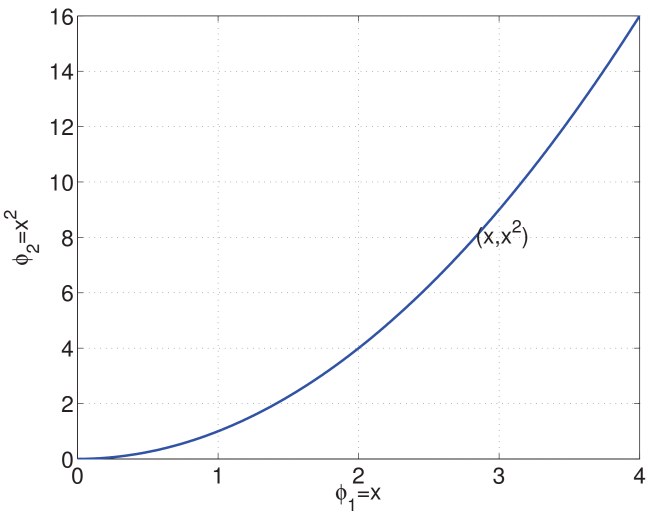

Figure 1: SVM Classification

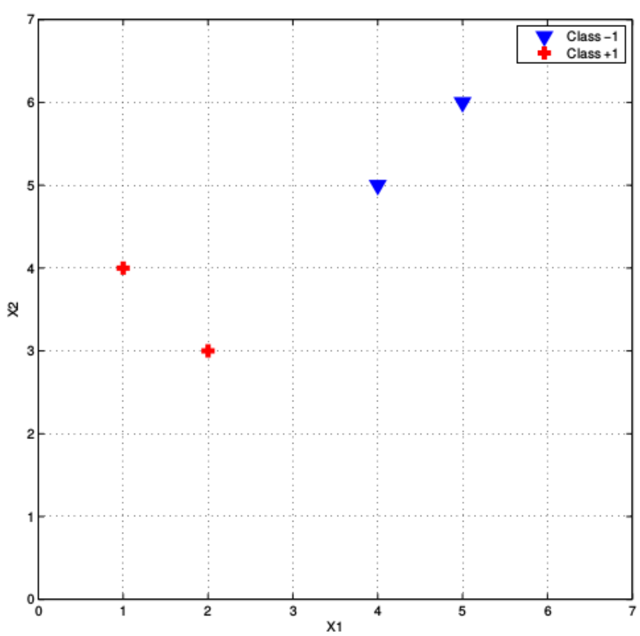

Figure 2: hard margin