**Practice Problems: Support Vector Machines**

1. Consider the following training data set of (seven) points $X$'s in a plane and their binary class label $y$'s:

| $X$ | (1, 0) | (0, 1) | (0, -1) | (-1, 0) | (0, 2) | (0, -2) | (-2, 0) |
|---|---|---|---|---|---|---|---|
| $y$ | -1 | -1 | -1 | +1 | +1 | +1 | +1 |

We perform the following non-linear transform of the input vector $X = (x_1, x_2)$ to obtain the transformed feature vector $Z = (z_1, z_2) = (\phi_1(X), \phi_2(X))$, with $\phi_1(X) = x_2^2 - 2x_1 + 3$, $\phi_2(X) = x_1^2 - 2x_2 - 3$. Write the equation of the optimal separating hyperplane in transformed space $Z$. Explain your answer.

2. A hypothetical two class hard margin linear SVM has the following values of Lagrange multipliers $\alpha$, support vectors, and output class labels $y$:

| $\alpha$ | *Support vector* | $y$ |
|---|---|---|
| 1 | (0, –1, 1) | +1 |
| 1 | (0, 2, –1) | –1 |
| 1 | (–1, 0, 2) | –1 |

Compute the predicted class label $y$ of this SVM when the input feature vector is (0.2, 0.8, 0.4).

3. Consider a set of 2-dimensional training data points ($x_1$, $x_2$) belonging to two classes +1 and -1 respectively as shown below. We design a linear hard margin SVM to classify them. The equation of the optimal separating hyperplane is?

+1: (3, 1), (3, −1), (6, 1), (6, −1)
-1: (1, 0), (0, 1), (0, −1), (−1, 0).

4. Consider the set of training data: Find the corresponding dual cost function (as a function of $\alpha_i$'s only)

| *Feature* | *label* | *feature* | *label* |
|---|---|---|---|
| $x_1$ = (0,0,0) | $d_1 = -1$ | $x_5$ = (1,0,0) | $d_5 = 1$ |
| $x_2$ = (0,0,1) | $d_2 = 1$ | $x_6$ = (1,0,1) | $d_6 = -1$ |
| $x_3$ = (0,1,0) | $d_3 = 1$ | $x_7$ = (1,1,0) | $d_7 = -1$ |
| $x_4$ = (0,1,1) | $d_4 = -1$ | $x_8$ = (1,1,1) | $d_8 = 1$ |