CS5350: Machine Learning	*Due 21 Oct 2008*

HW4: Loss and Optimization

1 Written Exercises

Answer the following questions in 25-100 words each:

1. Consider the degree-two polynomial kernel defined by $K(\boldsymbol{x}, \boldsymbol{z}) = (1 + \boldsymbol{x}^\top \boldsymbol{z})^2$. Expand this out completely for the three-dimensional case (i.e., $\boldsymbol{x} = \langle x_1, x_2, x_3 \rangle$ and $\boldsymbol{z} = \langle z_1, z_2, z_3 \rangle$. Verify that this has the same form as the quadratic expansion, although with different coefficients on the terms.

2. Continuing from the previous question, what is the form of Φ so that $K(\boldsymbol{x}, \boldsymbol{z}) = \Phi(\boldsymbol{x})^\top \Phi(\boldsymbol{z})$? (You need only consider the three-dimensional data case.) How does this differ from the expansion $\Phi(\boldsymbol{x}) = \langle x_1, x_2, x_3, x_1^2, x_2^2, x_3^2, x_1 x_2, x_1 x_3, x_2, x_3 \rangle$ that we discussed in class?

3. Consider optimizing an SVM with *squared* loss on the ξ variables. That is, an optimization problem of the form:

$$\min_{\boldsymbol{w}, b} \frac{1}{2} ||\boldsymbol{w}||^2 + \lambda \sum_n \xi_n^2$$

$$\text{s.t.} y_n(\boldsymbol{w}^\top \boldsymbol{x}_n + b) \geq 1 - \xi_n \qquad (\forall n)$$

$$\xi_n \geq 0 \qquad (\forall n)$$

Construct the dual formulation for this problem. In particular, construct the Lagrangian, optimize it with respect to $\boldsymbol{w}$ and b, plug these solutions back in and get an optimization problem just in terms of the dual (Lagrange) variables $\boldsymbol{\alpha}$. How does this compare to the dual formulation for the standard SVM?

4. (6350 only) For D dimensional data, consider using the degree d polynomial kernel defined by $K(\boldsymbol{x}, \boldsymbol{z}) = (1 + \boldsymbol{x}^\top \boldsymbol{z})^d$. What is the general form of the expansion? What are the coefficients on all the different forms in the expansion?