CS5350: Machine Learning *Due 30 Sep 2008*

HW4: Loss and Optimization

1 Written Exercises

Answer the following questions in 25-100 words each:

1. Suppose we are trying to find the minimum of the function $f(x) = 3x^2 - 2x + 1$, for univarite (scalar) x. First, find the minimum of this function using calculus. Second, verify that this function is convex (and therefore the point you found in the first step is a *global* minimum). Finally, perform (by hand – show your work) three steps of gradient descent with $\eta = 0.1$ and the initial point $x_0 = 1$. How close does it get to the true solution?

2. 0/1 loss is hard to optimize using gradient descent because it is poorly behaved (discontinuous, non-differentiable, etc.). A common "smooth" version of 0/1 loss is the *log-sigmoid* loss, which makes use of our favorite function, the sigmoid: $\sigma(z) = 1/(1+\exp[-z])$. In particular, we define our loss function $\ell(y, f(x)) = -\log(\sigma(yf(x)))$. First, verify that this is a reasonable loss function: when $f(x)$ is correct, the loss is low, and when $f(x)$ is incorrect, the loss is high. You may assume that $f(x)$ produces *real values*, where positive values mean class $+1$ and negative values mean class -1.

3. Consider optimizing a linear function under sigmoid loss, so that we have $f(\mathbf{X}, \boldsymbol{y}, \boldsymbol{w}) = \sum_n -\log(\sigma(y_n(\boldsymbol{w}^\top \boldsymbol{x}_n + b)))$. Compute the gradient of this function with respect to $\boldsymbol{w}$ and with respect to b so that we might construct a gradient descent algorithm.

4. (6350 only) Show that the loss function from the previous two questions is convex in both $\boldsymbol{w}$ and b.