| | |
|---|---|
| **Your Name** | |
| **Your UnetID** | |

## Error Minimization:

Given the loss function: $L = \sum_j^N [y_j - (w_o + w_1 x_j)]^2$, **calculate** $\frac{\partial L}{\partial w_0}$ **and** $\frac{\partial L}{\partial w_1}$

*The dark side of accuracy:*

**How many degrees of freedom can you expect to learn in any given model?**

We discussed degrees of freedom – Our example above has two weights, and therefore two degrees of freedom. The 15th order polynomial has 16 weights, and therefore 16 degrees of freedom. **Given that models with high degrees of freedom may overtrain, how does regularization help address this issue?**

The weight update for a perceptron is: $\boldsymbol{w \leftarrow w - \eta(y_i - h_w(x_i))x_i}$ , **How does the 'step direction' compare to the partial derivative found in part 1?**