DSC 140A - Discussion 02

Problem 1.

You are building a predictive model for a dataset representing measurements of houses. Each data point has three features: the number of bedrooms (x_1), the size in square feet (x_2), and the age of the house in years (x_3). The target is the house price (y).

The dataset is:

x_1	x_2	x_3	y
2	1200	20	250
3	1500	15	300
5	1700	10	400
1	800	25	200

We aim to fit a linear model of the form:

$$y = w_0 + w_1x_1 + w_2x_2 + w_3x_3$$

using stochastic gradient descent (SGD) and the following loss function for each data point:

$$L(H(\vec{x}^{(i)}; \vec{w}), y_i) = (y_i - Aug(\vec{x}^{(i)}) \cdot \vec{w})^2.$$

Perform three iterations of SGD and report the final values of w_0, w_1, w_2 and w_3. As a reminder, the steps to take are:

1. Derive the gradients of the loss function with respect to each parameter w_0, w_1, w_2 and w_3.
2. Perform stochastic gradient descent for three iterations with a batch size of 2 with the following settings:
   - Initial values: $w_0 = 0$, $w_1 = 0$, $w_2 = 0$, $w_3 = 0$
   - Learning rate: $\eta = 0.01$
3. Update the parameters step-by-step using the gradients.
4. Report the final values of w_0, w_1, w_2 and w_3.

Problem 2.

Suppose you are working on a machine learning model where underestimating the target variable (y) is penalized more than overestimating it. The custom loss function is defined as:

$$L(y, H(x)) = \begin{cases} 2(y - H(x))^2, & \text{if } y > H(x), \\ (y - H(x))^2, & \text{otherwise.} \end{cases}$$

The model predicts the target variable using the equation:

$$H(x) = w_0 + w_1x$$

where w_1 is the weight, w_0 is the bias, and x is the input feature.

a) Sketch a graph of the loss function $L(y, H(x))$ as a function of $H(x)$, for a fixed value of y.

b) Derive the gradients of the custom loss function with respect to the parameters w_1 and w_0.