CS 615 - Deep Learning

Assignment 3 - Learning and Basic Architectures

Winter 2024

Alec Peterson (ap3842@drexel.edu)

1 Theory

1. For the function $J = (x_1w_1 - 5x_2w_2 - 2)^2$, where $w = [w_1, w_2]^T$ are our weights to learn:

   (a) What are the partial gradients, $\frac{\partial J}{\partial w_1}$ and $\frac{\partial J}{\partial w_2}$? Show work to support your answer (6pts).

   i. $\frac{\partial J}{\partial w_1} = 2(x_1w_1 - 5x_2w_2 - 2)(x_1 - 0 - 0)$
   $= 2x_1^2w_1 - 10x_1x_2w_2 - 4x_1$

   ii. $\frac{\partial J}{\partial w_2} = 2(x_1w_1 - 5x_2w_2 - 2)(0 - 5x_2 - 0)$
   $= -10x_1x_2w_1 + 50x_2^2w_2 + 20x_2$

   (b) What are the value of the partial gradients, given current values of $w = [0, 0]^T, x = [1, 1]$ (4pts)?
   $\frac{\partial J}{\partial w_1} = 2(1)(0) - 10(1)(1)(0) - 4(1) = 0 - 0 - 4 = -4$

   $\frac{\partial J}{\partial w_2} = -10(1)(1)(0) + 50(1)(0) + 20(1) = 20$

2. Given the objective function $J = \frac{1}{4}(x_1w_1)^4 - \frac{4}{3}(x_1w_1)^3 + \frac{3}{2}(x_1w_1)^2$:

   (a) What is the gradient $\frac{\partial J}{\partial w_1}$ (2pts)?
   $\frac{\partial J}{\partial w_1} = \frac{4}{4}(x_1w_1)^3(x_1) - \frac{12}{3}(x_1w_1)^2(x_1) + \frac{6}{2}(x_1w_1)(x_1)$
   $= x_1^4w_1^3 - x_1^3w_1^2 + 3x_1^2w_1$

   (b) What are the locations of the extrema points for this objective function J if $x_1 = 1$? Recall that to find these you take the derivative of the objective function with respect to the unknown, set that equal to zero and solve for said unknown (in this case, w_1). (5pts)

   $\frac{\partial J}{\partial w_1} = 0 = (1)w_1^3 - 4(1)w_1^2 + 3(1)w_1$
   $= w_1^3 - 4w_1^2 + 3w_1 = w_1(w_1^2 - 4w_1 + 3)$
   $\implies 0 = w_1(w_1 - 1)(w_1 - 3)$
   $\implies w_1 = 0, w_1 = 1, w_1 = 3$

   (c) What does J evaluate to at each of your extrema points, again when $x_1 = 1$ (3pts)?

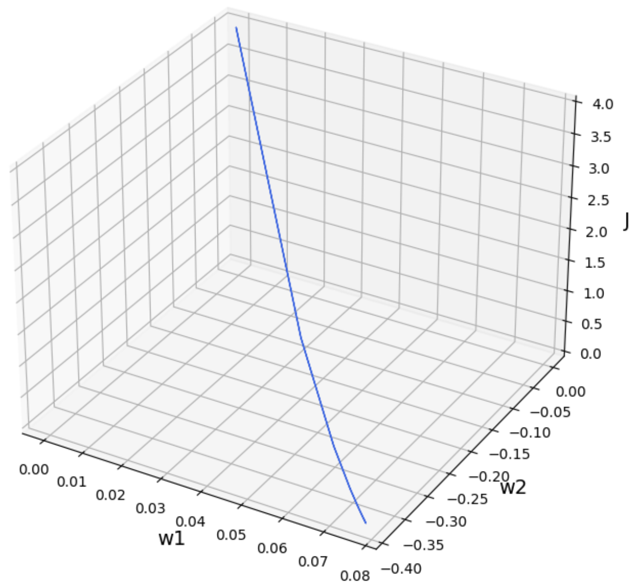
   i. $x_1 = 1, w_1 = 0$:
   $J = \frac{1}{4}(1 * 0)^4 - \frac{4}{3}(1 * 0)^3 + \frac{3}{2}(1 * 0)^2$
   $= 0$

   ii. $x_1 = 1, w_1 = 1$:
   $J = \frac{1}{4}(1 * 1)^4 - \frac{4}{3}(1 * 1)^3 + \frac{3}{2}(1 * 1)^2$
   $= \frac{5}{12}$

   iii. $x_1 = 1, w_1 = 3$:
   $J = \frac{1}{4}(1 * 3)^4 - \frac{4}{3}(1 * 3)^3 + \frac{3}{2}(1 * 3)^2$
   $= -\frac{9}{4}$

2 Visualizing Gradient Descent

See submitted code and below plot.

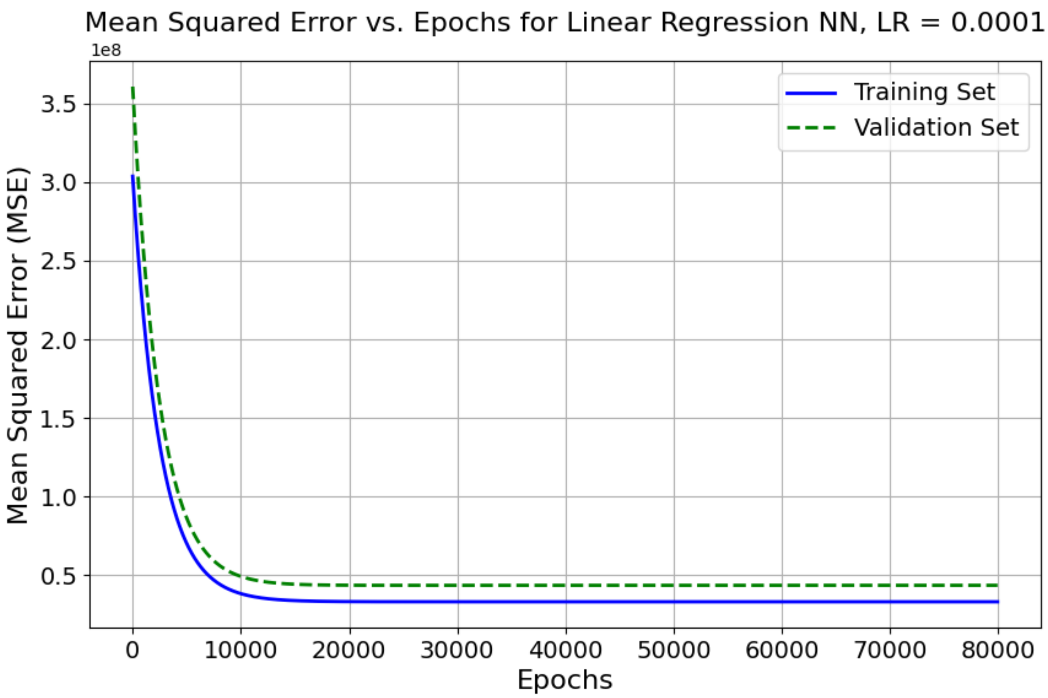


3 Updating Fully Connected Layer's Weights and Biases

See submitted code.

4 Linear Regression

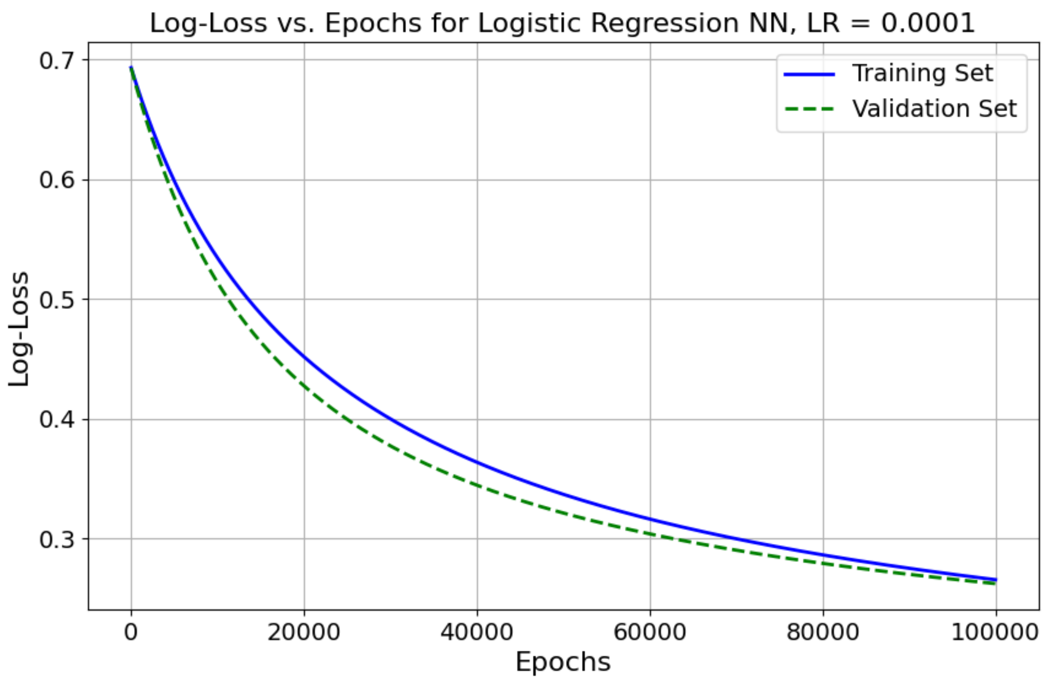
See submitted code, and below figures and reported error metric(s)



- Final Training RMSE = 5757.89
- Final Validation RMSE = 6604.31
- Final Training SMAPE = 0.1805
- Final Validation SMAPE = 0.1830

5 Logistic Regression

See submitted code, and below figures and reported error metric(s)



- Final Training Accuracy = 0.9286
- Final Validation Accuracy = 0.9067