Homework 3

David A. Halloran

February 8, 2022

Theory

1. $J = (x_1w_1 - 5x_2w_2 - 2)^2, w = \begin{bmatrix} w_1 & w_2 \end{bmatrix}^T$

   (a) Calculating the partial derivatives of w_1 and w_2. We do this one at a time by holding the x values and the other w value constant.
   Let's start with w_1. Apply the chain rule. So first bring down the exponent and chain it with the polynomial without the exponent. This step gives us

   $$2(x_1w_1 - 5x_2w_2 - 2) * \frac{\partial J}{\partial w_1}(x_1w_1 - 5x_2w_2 - 2)$$

   So next as indicated we now apply the partial derivative to the polynomial. With all else being constant it drops away to just x_1w_1. Taking the derivative of this with respect to w gives us $x_1 * 1w_1^0$. As anything raised to the 0th power is just one this leaves just x_1. So combining all of this we get a final answer of

   $$2(x_1w_1 - 5x_2w_2 - 2)x_1$$

   Now lets do w_2 It is the same process with just keeping w_1 constant this time. So we will use the chain rule again and chain together the polynomial with the exponent brought down multiplied by the partial of the polynomial. This again gives us the same general form

   $$2(x_1w_1 - 5x_2w_2 - 2) * \frac{\partial J}{\partial w_2}(x_1w_1 - 5x_2w_2 - 2)$$

   So when evaluating the second half the constants drop away giving us $-5x_2 * 1w_2^0$. This simplifies to $-5x_2$. Now we can take this and multiply the outer terms to simplify further so this gives us a final answer of:

   $$-10x_2(x_1w_1 - 5x_2w_2 - 2)$$

(b) Given w = [0, 0] and x = [1, 1], we can use these numbers to solve for the partial gradients derived in part 1a.
Let's do w1 first, remember the derivative of that one was

$$2(x_1w_1 - 5x_2w_2 - 2)x_1$$

So plugging in our figures the value will be:

$$2 * (1 * 0 - 5 * 1 * 0 - 2) * 1$$

$$2 * (0 - 0 - 2) * 1$$

$$2 * -2 * 1 = -4$$

For the partial of w_2 we had the following derivative formula:

$$-10x_2(x_1w_1 - 5x_2w_2 - 2)$$

This will reduce to:

$$-10 * 1 * (1 * 0 - 5 * 1 * 0 - 2)$$

$$-10 * (-2) = 20$$

2. $J = \frac{1}{4}(x_1w_1)^4 - \frac{4}{3}(x_1w_1)^3 + \frac{3}{2}(x_1w_1)^2$

(a) To calculate gradient of this function let us first distribute the outer terms to their respective inner terms. All combined this gives us

$$J = \frac{x^4w^4}{4} - \frac{4x^3w^3}{3} + \frac{3x^2w^2}{2}$$

Next we do the power rule on w while leave x constant. Conveniently for this example, each number pulled from the exponent is the same as the number in the denominator. So the power rule, the division, and keeping x constant gives us a final answer of the partial derivative of w as:

$$J = x^4w^3 - 4x^3w^2 + 3x^2w$$

(b) Lets find all the local extrema points of this function. We have a value of $x_1 = 1$ and we set the function equal to zero, then solve for w_2. So lets split up each section

$$0 = \frac{1}{4}(1 * w_1)^4$$

$$4 * 0 = 4 * \frac{1}{4}(1 * w_1)^4$$

$$0 = 4(1 * w_1)^4$$

$$0 = 4w_1^4$$

$$w_1 = 0$$

$$0 = \frac{4}{3}(1 * w_1)^3$$
$$3 * 0 = 3 * \frac{4}{3}(1 * w_1)^3$$
$$0 = 4w_1^3 3$$

$$2 * 0 = 2 * \frac{3}{2}(1 * w_1)^2$$
$$0 = 3w_1)^3$$

Plots

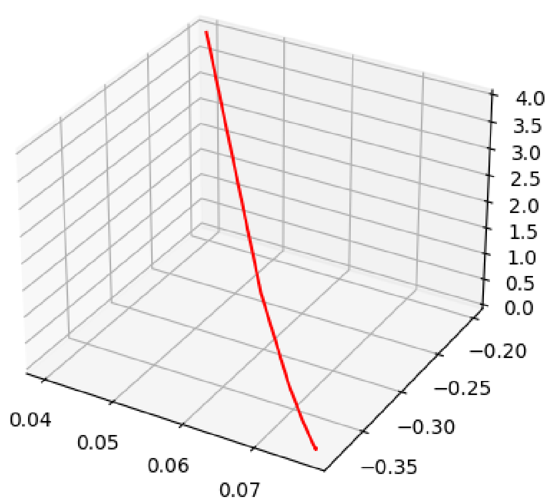

Figure 1: w1,w2, and J

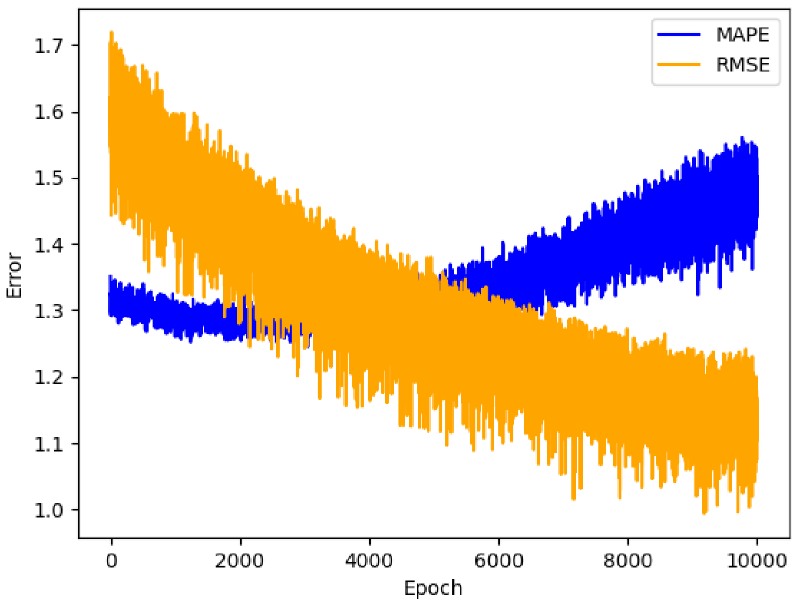

Figure 2: Training Data

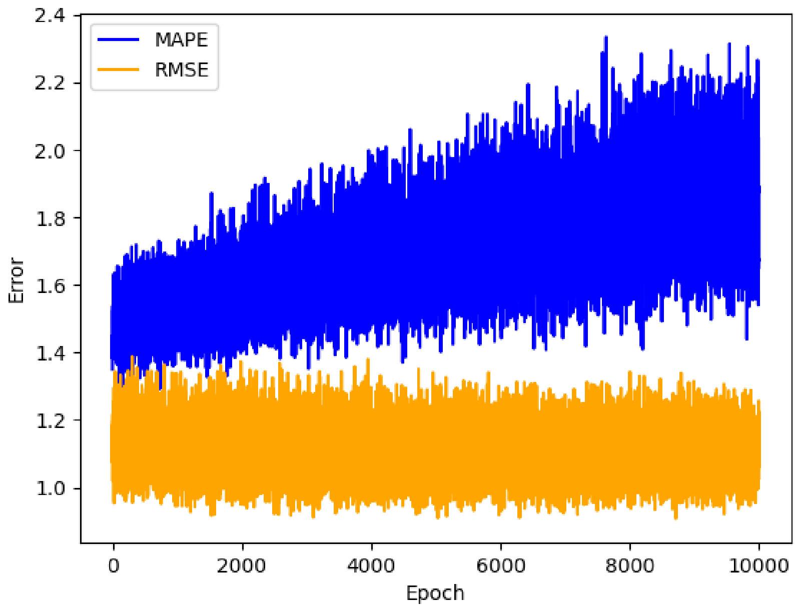

Figure 3: Validation Data