Assignment Solution: Electricity Consumption Prediction

Part (e): Line Search Methods

The Goal

In the previous steps, we calculated the direction of steepest descent (the negative gradient). Now, we must determine **how far** to move in that direction. This distance is called the *Step Size* (η).

We define a function $\phi(\eta)$ representing the Total Loss J for a specific step size η. Our goal is to find the η that minimizes this loss.

Setup:

- Current weights: $w^{(0)} = [2, 2]^T$
- Direction: $d = [-5, -21]^T$
- **Loss Function Formula:**

$$J(w_0, w_1) = \frac{1}{2}\left[(w_0 + 1w_1 - 3)^2 + (w_0 + 2w_1 - 5)^2 + (w_0 + 4w_1 - 9)^2 + (w_0 + 6w_1 - 13)^2 + (w_0 + 8w_1 - 17)^2\right]$$

The new weights for any given step size η are:

$$w(\eta) = w^{(0)} + \eta d = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \eta \begin{bmatrix} -5 \\ -21 \end{bmatrix} = \begin{bmatrix} 2 - 5\eta \\ 2 - 21\eta \end{bmatrix}$$

—

1. Binary Search Method

Logic: We check the midpoint m of our interval. To see if the true minimum is to the left or right, we check a point slightly to the right ($m + \epsilon$) to find the slope.

- If $\text{Loss}(m) < \text{Loss}(m + \epsilon)$: Slope is positive (uphill). Go **Left**.
- If $\text{Loss}(m) > \text{Loss}(m + \epsilon)$: Slope is negative (downhill). Go **Right**.

Iteration 1

Interval: $[0, 1]$. Midpoint $m = 0.5$.

Step A: Calculate Loss at Midpoint ($\eta = 0.5$) Weights:

$$w_0 = 2 - 5(0.5) = -0.5$$

$$w_1 = 2 - 21(0.5) = -8.5$$

Substitute into the full loss function:

$$J = \frac{1}{2}\left[(-0.5 + 1(-8.5) - 3)^2 + (-0.5 + 2(-8.5) - 5)^2 + (-0.5 + 4(-8.5) - 9)^2 + (-0.5 + 6(-8.5) - 13)^2 + (-0.5 + 8(-8\right.$$

$$J = \frac{1}{2}\left[(-0.5 - 8.5 - 3)^2 + (-0.5 - 17.0 - 5)^2 + (-0.5 - 34.0 - 9)^2 + (-0.5 - 51.0 - 13)^2 + (-0.5 - 68.0 - 17)^2\right]$$

$$J = \frac{1}{2}\left[(-12)^2 + (-22.5)^2 + (-43.5)^2 + (-64.5)^2 + (-85.5)^2\right]$$

$$J = \frac{1}{2}[144 + 506.25 + 1892.25 + 4160.25 + 7310.25]$$

$$J = \frac{1}{2}(14013) = \mathbf{7006.5}$$

Step B: Calculate Loss at Perturbed Point ($\eta = 0.501$) Weights:

$$w_0 = 2 - 5(0.501) = -0.505$$

$$w_1 = 2 - 21(0.501) = -8.521$$

Substitute into the full loss function:

$$J = \frac{1}{2}\left[(-0.505 + 1(-8.521) - 3)^2 + (-0.505 + 2(-8.521) - 5)^2 + (-0.505 + 4(-8.521) - 9)^2 + (-0.505 + 6(-8.521) - 1\right.$$

$$J \approx \frac{1}{2}\left[(-12.026)^2 + (-22.547)^2 + (-43.589)^2 + (-64.631)^2 + (-85.673)^2\right]$$

$$J \approx \frac{1}{2}[144.62 + 508.37 + 1900.00 + 4177.16 + 7339.86]$$

$$J \approx \frac{1}{2}(14070.01) = \mathbf{7035.0}$$

Decision: Since $7006.5 < 7035.0$, the function is increasing. We went too far. **New Interval:** $[0, 0.5]$.

Iteration 2

Interval: $[0, 0.5]$. Midpoint $m = 0.25$.

Step A: Calculate Loss at Midpoint ($\eta = 0.25$) Weights:

$$w_0 = 2 - 5(0.25) = 0.75$$

$$w_1 = 2 - 21(0.25) = -3.25$$

Substitute into the full loss function:

$$J = \frac{1}{2}\left[(0.75 + 1(-3.25) - 3)^2 + (0.75 + 2(-3.25) - 5)^2 + (0.75 + 4(-3.25) - 9)^2 + (0.75 + 6(-3.25) - 13)^2 + (0.75 + 8(-\right.$$

$$J = \frac{1}{2}\left[(-5.5)^2 + (-10.75)^2 + (-21.25)^2 + (-31.75)^2 + (-42.25)^2\right]$$

$$J = \frac{1}{2}[30.25 + 115.56 + 451.56 + 1008.06 + 1785.06]$$

$$J = \frac{1}{2}(3390.5) = \mathbf{1695.25}$$

Step B: Calculate Loss at Perturbed Point ($\eta = 0.251$) Weights:

$$w_0 = 2 - 5(0.251) = 0.745$$

$$w_1 = 2 - 21(0.251) = -3.271$$

Substitute into the full loss function:

$$J = \frac{1}{2}\left[(0.745 + 1(-3.271) - 3)^2 + (0.745 + 2(-3.271) - 5)^2 + (0.745 + 4(-3.271) - 9)^2 + (0.745 + 6(-3.271) - 13)^2 + (0\right.$$

$$J \approx \frac{1}{2}\left[(-5.526)^2 + (-10.797)^2 + (-21.339)^2 + (-31.881)^2 + (-42.423)^2\right]$$

$$J \approx \frac{1}{2}[30.54 + 116.58 + 455.35 + 1016.39 + 1799.71]$$

$$J \approx \frac{1}{2}(3418.57) = \mathbf{1709.28}$$

Decision: Since $1695.25 < 1709.28$, the function is increasing. We are still too far to the right. **New Interval:** $[0, 0.25]$.

—

2. Golden-Section Search

Logic: We evaluate two interior points, M_1 (at 25% distance) and M_2 (at 75% distance). We keep the section containing the lower loss.

Iteration 1

Interval: $[0, 1]$. Points: $M_1 = 0.25$, $M_2 = 0.75$.

Step A: Evaluate M_1 ($\eta = 0.25$) (We calculated this exactly in Binary Search Iteration 2 Step A).

$$J(M_1) = \mathbf{1695.25}$$

Step B: Evaluate M_2 ($\eta = 0.75$) Weights:

$$w_0 = 2 - 5(0.75) = -1.75$$

$$w_1 = 2 - 21(0.75) = -13.75$$

Substitute into the full loss function:

$$J = \frac{1}{2}\left[(-1.75 + 1(-13.75) - 3)^2 + (-1.75 + 2(-13.75) - 5)^2 + (-1.75 + 4(-13.75) - 9)^2 + (-1.75 + 6(-13.75) - 13)^2 + \right.$$

$$J = \frac{1}{2}\left[(-18.5)^2 + (-34.25)^2 + (-65.75)^2 + (-97.25)^2 + (-128.75)^2\right]$$

$$J = \frac{1}{2}[342.25 + 1173.06 + 4323.06 + 9457.56 + 16576.56]$$

$$J = \frac{1}{2}(31872.5) = \mathbf{15936.25}$$

Decision: $J(M_1) < J(M_2)$ $(1695.25 < 15936.25)$. The minimum is in the left section. **New Interval:** $[0, 0.75]$.

Iteration 2

Interval: $[0, 0.75]$. New Points: $M_1 = 0 + 0.25(0.75) = 0.1875$. $M_2 = 0 + 0.75(0.75) = 0.5625$.

Step A: Evaluate M_1 ($\eta = 0.1875$) Weights:

$$w_0 = 2 - 5(0.1875) = 1.0625$$

$$w_1 = 2 - 21(0.1875) = -1.9375$$

Substitute into the full loss function:

$$J = \frac{1}{2}\left[(1.06 - 1.94 - 3)^2 + (1.06 - 3.88 - 5)^2 + (1.06 - 7.75 - 9)^2 + (1.06 - 11.62 - 13)^2 + (1.06 - 15.5 - 17)^2\right]$$

$$J = \frac{1}{2}\left[(-3.875)^2 + (-7.8125)^2 + (-15.6875)^2 + (-23.5625)^2 + (-31.4375)^2\right]$$

$$J = \frac{1}{2}[15.01 + 61.03 + 246.10 + 555.19 + 988.31]$$

$$J = \frac{1}{2}(1865.64) = \mathbf{932.82}$$

Step B: Evaluate M_2 ($\eta = 0.5625$) Weights:

$$w_0 = 2 - 5(0.5625) = -0.8125$$

$$w_1 = 2 - 21(0.5625) = -9.8125$$

Substitute into the full loss function:

$$J = \frac{1}{2}\left[(-0.81 - 9.81 - 3)^2 + (-0.81 - 19.62 - 5)^2 + (-0.81 - 39.25 - 9)^2 + (-0.81 - 58.88 - 13)^2 + (-0.81 - 78.5 - 17\right.$$

$$J = \frac{1}{2}\left[(-13.625)^2 + (-25.4375)^2 + (-49.0625)^2 + (-72.6875)^2 + (-96.3125)^2\right]$$

$$J = \frac{1}{2}[185.64 + 647.06 + 2407.13 + 5283.47 + 9276.10]$$

$$J = \frac{1}{2}(17799.4) = \mathbf{8899.7}$$

Decision: $J(M_1) < J(M_2)$ $(932.82 < 8899.7)$. The minimum is in the left section. **New Interval:** $[0, 0.5625]$.