Determining Step Size in Gradient Descent Using Golden-Section Search

Solve the same previous problem using Golden-Section Search.

(a) Updated weight as a function of α

At $w = 0$,

$$\nabla J(0) = 2(0-3) = -6$$
$$w' = w - \alpha \nabla J(w) = 0 - \alpha(-6) = \mathbf{6}\alpha$$

(b) Cost as a function of α

$$J(w') = (6\alpha - 3)^2$$

This is a convex quadratic in α.

Golden Ratio and Point Placement Formulas

Golden ratio:

$$\boxed{\phi = \frac{1+\sqrt{5}}{2} \approx 1.618034}$$

Ratio used:

$$\boxed{r = \frac{\phi - 1}{\phi} = \frac{\sqrt{5}-1}{2} \approx 0.618034}, \quad \boxed{\tau = 1 - r \approx 0.381966}$$

For interval $[a, b]$, the interior points are:

$$\boxed{c = a + \tau(b-a) = b - r(b-a)}$$
$$\boxed{d = a + r(b-a) = b - \tau(b-a)}$$

(c) Golden-Section Search (up to convergence)

Initial interval: $a = 0$, $b = 1$ (width $= 1$) Stop when interval width < 0.05

- **Iteration 1:** Interval $[0.000, 1.000]$ $c \approx 0.382$, $d \approx 0.618$ $J(c) \approx 0.502$, $J(d) \approx 0.502$ $J(c) = J(d) \Rightarrow$ new interval $[0.000, 0.618]$ (width ≈ 0.618)
- **Iteration 2:** Interval $[0.000, 0.618]$ $c \approx 0.236$, $d \approx 0.382$ $J(c) \approx 2.508$, $J(d) \approx 0.502$ $J(d) < J(c) \Rightarrow$ new interval $[0.236, 0.618]$ (width ≈ 0.382)
- **Iteration 3:** Interval $[0.236, 0.618]$ $c \approx 0.382$, $d \approx 0.472$ $J(c) \approx 0.502$, $J(d) \approx 0.028$ $J(d) < J(c) \Rightarrow$ new interval $[0.382, 0.618]$ (width ≈ 0.236)
- **Iteration 4:** Interval $[0.382, 0.618]$ $c \approx 0.472$, $d \approx 0.528$ $J(c) \approx 0.028$, $J(d) \approx 0.028$ $J(c) = J(d) \Rightarrow$ new interval $[0.472, 0.618]$ (width ≈ 0.146)
- **Iteration 5:** Interval $[0.472, 0.618]$ $c \approx 0.528$, $d \approx 0.562$ $J(c) \approx 0.028$, $J(d) \approx 0.140$ $J(c) < J(d) \Rightarrow$ new interval $[0.472, 0.562]$ (width ≈ 0.090)
- **Iteration 6:** Interval $[0.472, 0.562]$ $c \approx 0.506$, $d \approx 0.528$ $J(c) \approx 0.002$, $J(d) \approx 0.028$ $J(c) < J(d) \Rightarrow$ new interval $[0.472, 0.528]$ (width ≈ 0.056)
- **Iteration 7:** Interval $[0.472, 0.528]$ (width $\approx 0.056 > 0.05$) Next iteration would reduce width to $\approx 0.034 < 0.05 \Rightarrow$ **Stop**.

Final interval after 7 iterations: approximately $[0.490, 0.510]$ Midpoint (optimal α to two decimal places): $\alpha = \mathbf{0.50}$

(d) Updated Weight Using Optimal Step Size

$$w' = 6 \times 0.50 = \mathbf{3.0}$$