
Artificial Intelligence Lab 7 : Gradient Descent in 1-dimension

Let $L: \mathbb{R} \rightarrow \mathbb{R}$ be a function, and let:

- (1) $L_* = \min_{w \in \mathbb{R}} L(w)$ be its **minimum** value
- (2) $w_* = \arg \min_{w \in \mathbb{R}} L(w)$ be the **minimiser**, i.e., the argument for which the minimum value is attained (notice $L_* = L(w_*)$).

The gradient descent algorithm to find w_* of a function $L(w)$ is given as follows:

$$w_{t+1} = w_t - \alpha \frac{dL(w)}{dw} \Big|_{w=w_t}, \quad (1)$$

where

- (i) $\alpha > 0$ is a step-size parameter
- (ii) $\frac{dL(w)}{dw} \Big|_{w=w_t}$ is the derivative of $L(w)$ with respect to w evaluated at $w = w_t$.

Let $L(w) = \frac{a}{2}w^2 + bw + c$.

[Q1] [50 Marks] Plot the function $L(w)$ (in red) $L_* = L(w_*)$ (blue dot) and w_* (green star), for the following cases:

- (1) $a = 1, b = 0, c = 0$.
- (2) $a = 0.1, b = 0, c = 0$
- (3) $a = 10, b = 0, c = 0$
- (4) $a = 1, b = 0, c = 10$
- (5) $a = 1, b = 1, c = 1$
- (6) $a = 0.1, b = -1, c = -1$

Calculate L_* and w_* with pen and paper. Choose appropriate axis range so that L_* and w_* can be displayed in a proper manner.

[Q2] [50 Marks] Run the gradient descent algorithm for $t = 1, \dots, 10$ iterations. For each iteration display

- (1) The complete function $L(w)$ (for w in an appropriate range used in [Q1]), and $L(w_t)$ in one plot
- (2) $w_t - w_*$ in a different plot.

Run the gradient descent algorithm for various values of α such that : (i) the iterates w_t are on the same side of w_* and converge to w_* , (ii) the iterates w_t oscillate on both sides of w_* and converge to w_* , (iii) the iterates w_t oscillate on both sides of w_* , but diverge to ∞ . Please derive these cases with pen and paper first and then proceed to code.