DA311 Machine Learning Lab

Assignment 1

Date: August 8th, 2023

1. Generate a set of points around a line $y = ax + b$
   (a) Choose $a = 2$ and $b = 3$.
   (b) Select the range for x as $[-10, 10]$ and generate $n = 100$ values for x in that interval.
   (c) Compute the values of y for each x as $y_i = 2x_i + 3$.
   (d) Plot the line $y = 2x + 3$ in black color.
   (e) Generate a set of n points around the line using the equation,

$$y_i = 2x_i + 3 + \mathcal{N}(0, 1) \tag{1}$$

   where, $\mathcal{N}(0, 1)$ is the zero-mean unity-variance normal distribution.
   (f) Show the scatter plot of these noisy points (in red color) on the same graph generated in step (d).

2. Plot the average error surface E for different values of a and b in the interval of $[-10 : 0.1 : 10]$.
   (a) Vary both a and b in steps of 0.1 in the interval $[-10, 10]$.
   (b) Compute the element-wise error as,

$$e_i = y_i - \hat{y}_i \tag{2}$$

   where, $\hat{y}_i = ax_i + b$ and y_i is computed using equation (1).
   (c) Compute the average error as,

$$E = \frac{1}{n} \sum_{i=1}^{n} {e_i}^2 \tag{3}$$

   (d) Compute the average error values for all combinations of a and b.
   (e) Plot the error surface with the values of a along x-axis, that of b along y-axis and E along z-axis.

3. Solve for a and b using Pseudo-inverse based approach on the points generated in question 1.

4. Solve for a and b using the Gradient Descent approach where, the values of $\mathbf{p} = (a, b)^T$ in the $(k + 1)^{\text{th}}$ iteration is updated as,

$$\mathbf{p}_{k+1} = \mathbf{p}_k - \eta \, \triangle_p \, E|_{p=p_k} \tag{4}$$

   Vary the update rate η and the initial values (a_0, b_0) and note the final solution after 100 iterations. Plot the trajectory of the solutions (a_k, b_k) for varying (a_0, b_0, η) on the contour plot of E on (a, b) plane.