Problem Set 1 Supervised Learning

DS542 DL4DS

Spring, 2025

Sicheng Yi (Tiger Yi)

Note: Refer to the equations in the *Understanding Deep Learning* textbook to solve the following problems.

Problem 2.1

To walk downhill on the loss function (equation 2.5), we measure its gradient with respect to the parameters ϕ_0 and ϕ_1. Calculate expressions for the slopes $\frac{\partial L}{\partial \phi_0}$ and $\frac{\partial L}{\partial \phi_1}$.

$$L[\phi] = \sum_{i=1}^{I} (f[x_i, \phi] - y_i)^2$$

$$L[\phi] = \sum_{i=1}^{I} (\phi_0 + \phi_1 x_i - y_i)^2$$

$$L[\phi] = \sum_{i=1}^{I} (\phi_0 + \phi_1 x_i - y_i)(\phi_0 + \phi_1 x_i - y_i)$$

$$L[\phi] = \sum_{i=1}^{I} (\phi_0^2 + \phi_0\phi_1 x_i - \phi_0 y_i + \phi_0\phi_1 x_i + \phi_1^2 x_i^2 - \phi_1 x_i y_i - \phi_0 y_i - \phi_1 x_i y_i + y_i^2)$$

$$L[\phi] = \sum_{i=1}^{I} (\phi_0^2 + 2\phi_0\phi_1 x_i - 2\phi_0 y_i + \phi_1^2 x_i^2 - 2\phi_1 x_i y_i + y_i^2)$$

$$\frac{\partial L}{\partial \phi_0} = \sum_{i=1}^{I} (2\phi_0 + 2\phi_1 x_i - 2y_i)$$

$$\frac{\partial L}{\partial \phi_1} = \sum_{i=1}^{I} (2\phi_1 x_i^2 - 2x_i y_i)$$

Problem 2.2

Show that we can find the minimum of the loss function in closed-form by setting the expression for the derivatives from Problem 2.1 to zero and solving for ϕ_0 and ϕ_1.

$$\frac{\partial L}{\partial \phi_0} = \sum_{i=1}^{I}(2\phi_0 + 2\phi_1 x_i - 2y_i) = 0$$

$$\frac{\partial L}{\partial \phi_1} = \sum_{i=1}^{I}(2\phi_1 x_i^2 - 2x_i y_i) = 0$$

For Equation 2:

$$\sum_{i=1}^{I} 2\phi_1 x_i^2 - \sum_{i=1}^{I} 2x_i y_i = 0$$

$$2\phi_1 \sum_{i=1}^{I} x_i^2 = 2\sum_{i=1}^{I} x_i y_i$$

$$\phi_1 = \frac{\sum_{i=1}^{I} x_i y_i}{\sum_{i=1}^{I} x_i^2}$$

For Equation 1:

$$\sum_{i=1}^{I} 2\phi_0 + \sum_{i=1}^{I} 2\phi_1 x_i - \sum_{i=1}^{I} 2y_i = 0$$

$$2I\phi_0 + 2\phi_1 \sum_{i=1}^{I} x_i - 2\sum_{i=1}^{I} y_i = 0$$

$$I\phi_0 = \sum_{i=1}^{I} y_i - \phi_1 \sum_{i=1}^{I} x_i$$

$$\phi_0 = \frac{1}{I}(\sum_{i=1}^{I} y_i - \phi_1 \sum_{i=1}^{I} x_i)$$

$$\phi_0 = \frac{1}{I}(\sum_{i=1}^{I} y_i - \frac{\sum_{i=1}^{I} x_i y_i \cdot \sum_{i=1}^{I} x_i}{\sum_{i=1}^{I} x_i^2})$$