Problem Set 1 – Supervised Learning

DS542 – DL4DS

Spring, 2025

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 $\frac{\partial L}{\partial \phi_1}$.

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

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

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

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.

To find the minimum of the loss function, we set the partial derivatives to zero and solve for ϕ_0 and ϕ_1.

Setting the derivatives to zero:

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

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

Simplifying the first equation:

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

Simplifying the second equation:

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

Solving these simultaneous equations gives:

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

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