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.
Disclaimer: Used GPT for formatting latex of derivation. Work was done on paper and sent to GPT for latex transcription.

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

Equation 2.5:

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

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

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

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

And derivation for the other parameter:

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

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

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.

Need to set the derivatives equal to 0 and solve.

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

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

And for the other parameter:

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

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

$$(\bar{y} - \phi_1 \bar{x}) \sum_{i=1}^{I} x_i + \phi_1 \sum_{i=1}^{I} x_i^2 = \sum_{i=1}^{I} x_i y_i$$

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

$$\phi_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$