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

$$L[\phi] = \sum_{i=1}^{I} (\mathrm{f}[x_i, \phi] - y_i)^2 = \sum_{i=1}^{I} (\phi_0 + \phi_1 x_i - y_i)^2 . \quad (2.5)$$

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

$$\boxed{2\sum_{i=1}^{I} (\phi_0 + \phi_1 x_i - y_i) \underbrace{\frac{d}{d\phi}\phi_0}_{1}} = 2\sum_{i=1}^{I} (\phi_0 + \phi_1 x_i - y_i) \cdot 1 \quad ※$$

= chain rule

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

$$= 2\sum_{i=1}^{I} (\phi_0 + \phi_1 x_i - y_i) \cdot \underbrace{\frac{d}{d\phi}\phi_1 x_i}_{x_i} = 2\sum_{i=1}^{I} (\phi_0 + \phi_1 x_i - y_i) \cdot x_i \quad ※$$

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.

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

$\div 2$

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

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

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

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

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

$$\phi_0 = \bar{y} - \bar{x} \quad ※$$