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

Answer:

1.

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

2.

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

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.
Answer:

1.

$$
\begin{aligned}
\frac{\partial L}{\partial \phi_0} &= 0 \\
\sum_{i=1}^{I} 2(\phi_0 + \phi_1 x_i - y_i) &= 0 \\
\sum_{i=1}^{I} (\phi_0 + \phi_1 x_i - y_i) &= 0 \\
\sum_{i=1}^{I} (\phi_0) + \sum_{i=1}^{I} (\phi_1 x_i) - \sum_{i=1}^{I} (y_i) &= 0 \\
\sum_{i=1}^{I} (\phi_0) &= \sum_{i=1}^{I} (y_i) - \sum_{i=1}^{I} (\phi_1 x_i) \\
I(\phi_0) &= \sum_{i=1}^{I} (y_i) - \sum_{i=1}^{I} (\phi_1 x_i) \\
\phi_0 &= \frac{\sum_{i=1}^{I} (y_i)}{I} - \phi_1 \frac{\sum_{i=1}^{I} x_i}{I} \\
\phi_0 &= \overline{y} - \phi_1 \overline{x}
\end{aligned}
$$

2.

$$\frac{\partial L}{\partial \phi_1} = 0$$

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

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

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

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

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

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

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

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

$$\phi_1 \sum_{i=1}^{I} x_i^2 - I \phi_1 \overline{x}^2 = \sum_{i=1}^{I} y_i x_i - \overline{y} I \overline{x}$$

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

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

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