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

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

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

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

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.

$$2 \sum_{i=1}^{I} (\tilde{\phi}_0 + \tilde{\phi}_1 x_i - y_i) \stackrel{!}{=} 0$$

$$\sum_{i=1}^{I} (\tilde{\phi}_0 + \tilde{\phi}_1 x_i - y_i) = n \cdot \tilde{\phi}_0 + \tilde{\phi}_1 \sum_{i=1}^{I} x_i - \sum_{i=1}^{I} y_i \Rightarrow$$

$$\Rightarrow \tilde{\phi}_0 = \frac{\sum y_i - \tilde{\phi}_1 \sum x_i}{n} = \frac{1}{n} \sum y_i - \frac{1}{n} \tilde{\phi}_1 \sum x_i$$

$$\tilde{\phi}_0 = \bar{y} - \tilde{\phi}_1 \bar{x}$$

mean of all x_i

mean of y_i

$$2\sum_{i=1}^{I}(\tilde{\phi}_0+\tilde{\phi}_1 x_i - y_i)x_i \overset{!}{=} 0 \Rightarrow \sum_{i=1}^{I}(x_i\tilde{\phi}_0+\tilde{\phi}_1(x_i)^2 - y_i x_i) = 0 \Rightarrow \tilde{\phi}_0 \sum x_i + \tilde{\phi}_1 \sum x_i^2 = \sum x_i y_i$$

$$\tilde{\phi}_0 \sum x_i + \tilde{\phi}_1 \sum x_i^2 = \sum x_i y_i$$

$$(\bar{y} - \tilde{\phi}_1 \bar{x}) \sum x_i + \tilde{\phi}_1 \sum x_i^2 = \sum x_i y_i$$

$$\bar{y} \sum x_i - \tilde{\phi}_1 \bar{x} \sum x_i + \tilde{\phi}_1 \sum x_i^2 = \sum x_i y_i$$

$$\tilde{\phi}_1 \left(\sum x_i^2 - \bar{x} \sum x_i\right) = \sum x_i y_i - \bar{y} \sum x_i$$

$$\tilde{\phi}_1 = \frac{\sum x_i y_i - \bar{y}\sum x_i}{\sum x_i^2 - \bar{x}\sum x_i}$$

$\sum x_i y_i - \bar{y}\sum x_i \longrightarrow$ equal to $\sum (y_i - \bar{y})(x_i - \bar{x})$

$\sum x_i^2 - \bar{x}\sum x_i \longrightarrow$ equal to $\sum (x_i - \bar{x})^2$

$$\tilde{\phi}_1 = \frac{\sum_{i=1}^{I}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{I}(x_i - \bar{x})^2} = \frac{SXY}{SXX}$$

Sum of XY

Sum of X^2