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

Assume $e_i = (\phi_0 + \phi_1 x_i) - y_i$, so $L = \sum_{i=1}^{I} e_1^2$

$\frac{\partial}{\partial \phi_0} = \frac{\partial}{\partial e_i} \frac{\partial}{\partial \phi_1} = x_i \frac{\partial}{\partial e_i}$

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

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

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.

For ϕ_0 :

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

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

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

For ϕ_1 :

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

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

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