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

Solution:

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

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

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

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. Note that this works for linear regression but not for more complex models; this is why we use iterative model fitting methods like gradient descent (figure 2.4).

Solution:

$$\frac{\partial L}{\partial \phi_0} = 0, \frac{\partial L}{\partial \phi_0} = 0 \Rightarrow \begin{cases} I\phi_0 + \phi_1 \sum_{i=1}^{I} x_i - \sum_{i=1}^{I} y_i = 0 \\ \phi_0 \sum_{i=1}^{I} x_i + \phi_1 \sum_{i=1}^{I} x_i^2 - \sum_{i=1}^{I} y_i x_i = 0 \end{cases}$$

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

$$\phi_0 = \bar{y} - \phi_1 \bar{x}$$

$$(\bar{x} = \frac{1}{I}\sum_{i=1}^{I} x_i, \bar{y} = \frac{1}{I}\sum_{i=1}^{I} y_i)$$

Problem 2.3*

Consider reformulating linear regression as a generative model, so we have $g[y, \phi] = \phi_0 + \phi_1 y$. What is the new loss function? Find an expression for the inverse function $y = g^{-1}[x, \phi]$ that we would use to perform inference. Will this model make the same predictions as the discriminative version for a given training dataset $\{(x_i, y_i)\}$? One way to establish this is to write code that fits a line to three data points using both methods and see if the result is the same.

Solution:

See official solution for answer