Problem 7.1

May 11, 2024

1 Problem

A two-layer network with two hidden units in each layer can be defined as:

$$\begin{aligned} y = \phi_0 + \phi_1 a\left[\nu_{01} + \nu_{11} a[\theta_{01} + \theta_{11} x] + \nu_{21} a[\theta_{02} + \theta_{12} x]\right] \\ + \phi_2 a\left[\nu_{02} + \nu_{12} a[\theta_{01} + \theta_{11} x] + \nu_{22} a[\theta_{02} + \theta_{12} x]\right] \end{aligned}$$

Compute the derivatives of the output y with respect to each of the 13 parameters directly (not using back propagation). Use $\mathbb{I}$ for $\frac{\partial a[z]}{\partial z}$.

2 Answer