Deep Learning

Theoretical Exercises – Week 7 – Chapter 6

Exercises on the book "Deep Learning" written by Ian Goodfellow,
Yoshua Bengio, and Aaron Courville.
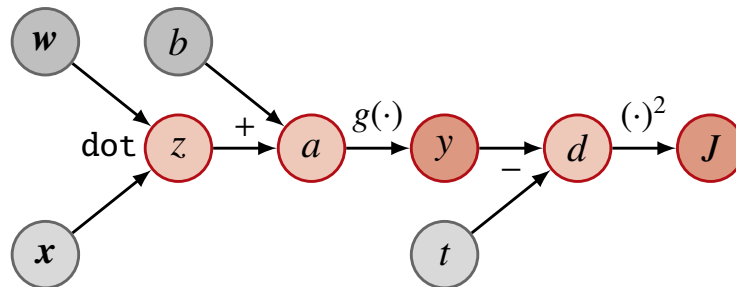Exercises and solutions by T. Méndez and G. Schuster

FS 2024

1 Exercises on Deep Feedforward Networks

1. What is the advantage of several hidden layers over one very large hidden layer?

2. Given is the following computational graph, where $x = [x_1\ x_2]^T$, $w = [w_1\ w_2]^T$ and

$$g(a) = \text{softplus(a)} = \log(1 + \exp(a)).$$

Calculate the following partial derivatives or gradients, respectively,

$$\frac{\partial J}{\partial b}, \qquad \nabla_w J, \qquad \nabla_x J, \qquad \text{and} \qquad \frac{\partial J}{\partial t}$$

by using the chain rule and the computational graph.