Deep Learning

Theoretical Exercises – Week 11 – Chapter 8

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 Optimization for Training Deep Models

1. Given is the following loss function

$$f(\boldsymbol{w}) = \frac{1}{4}\,w_1^4 + w_1^3 - \frac{17}{4}\,w_1^2 - 6\,w_1 + \frac{1}{5}\,w_2^4 + \frac{6}{5}\,w_2^3 + 89,$$

which has a global minimum at the point $\boldsymbol{c}_0 = [\,-4.572 \quad -4.5\,]^T$ and a local minimum at the point $\boldsymbol{c}_1 = [\,2.175 \quad -4.5\,]^T$.

(a) Determine the gradient of the loss function.

(b) Search for the global minimum by using gradient descent. In doing so, start at the point $\boldsymbol{w}^{(0)} = [4\ 4]^T$ and use the learning rate $\epsilon = 0.05$. Finish the learning algorithm after 10 iterations and check whether you have found the global minimum or not.

(c) Repeat Exercise (b) with the method of momentum and use $\alpha = 0.5$.