Optimization Algorithms

Applied ML in Engineering - Exercise 10

TU Berlin, Summer Term 2023

Prof. Dr.-Ing. Merten Stender – merten.stender@tu-berlin.de

Students are asked to implement different optimization algorithms and test their performance on a non-convex objective function $f(x) = 0.3 * x^4 - 0.1x^3 - 2x^2 - 0.8x$ as shown in Figure 1. Choose $x = -2.8$ as initial point for all optimizers.

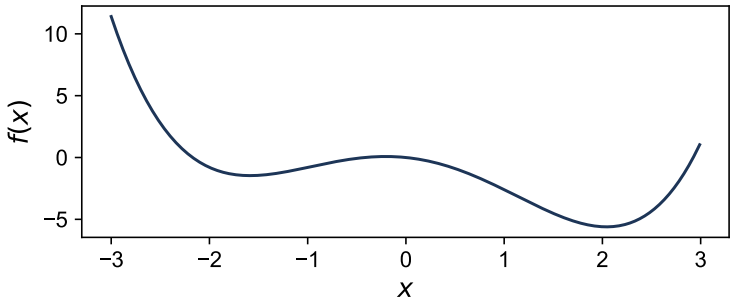

Figure 1: Objective function

Problem 1: Gradient Descent with Momentum

(a) Extent the provided implementation of gradient descent (`gradient_descent.py`) and include the momentum term as shown in the lecture

(b) Test how different values for the hyperparameter β will affect the optimization result and convergence when tested on the non-convex objective function.

(c) Find the optimal value for β that achieves the fastest convergence to the global minimum for the given function

Problem 2: Adagrad

(a) Extent the existing implementations by an adaptive learning rate as proposed in the Adagrad algorithm

(b) Test how Adagrad performs on the given sample function and compare against gradient descent and the momentum extension.

Problem 3: adam

(a) Extent the existing implementations by an adaptive learning rate and moment as proposed in the adam algorithm

(b) Test how adam performs on the given sample function and compare against gradient descent and Adagrad.