

Friedrich-Alexander-Universität Erlangen-Nürnberg
CML: Control, Machine Learning, and Numerics
Assignment 4

Due Date: 6:00pm, 9 July, 2023

1. In the lecture, we considered the set of points given below which shows labeled data: 5 points are in category A, indicated by circles and labelled by $y = (1, 0)^T$, and the rest 5 points are in category B, indicated by crosses and labelled by $y = (0, 1)^T$.

$$x_1 = [0.1, 0.3, 0.1, 0.6, 0.4, 0.6, 0.5, 0.9, 0.4, 0.7];$$

$$x_2 = [0.1, 0.4, 0.5, 0.9, 0.2, 0.3, 0.6, 0.2, 0.4, 0.6];$$

$$y = [\text{ones}(1, 5), \text{zeros}(1, 5); \text{zeros}(1, 5), \text{ones}(1, 5)];$$

We trained a neural network by standard Stochastic Gradient Descent (SGD) method to categorize a newly unseen point in \mathbb{R}^2 , i.e. construct a mapping that takes any point in \mathbb{R}^2 and returns either a circle or a cross. The details can be referred to the file "classification_NN.m". Next, please train the neural networks using

- (i) SGD with momentum
- (ii) Nesterov accelerated SGD (NAG)

and compare them with the standard SGD with learning rate 0.05.

For algorithms (i)-(ii), please select the parameters carefully such that they converge at least as fast as the standard SGD with learning rate 0.05. You should report the selected parameters (use the same notations as those in the slides) for each algorithm and the comparison results with the standard SGD with learning rate 0.05. (To report the comparison result, please present a figure to indicate the objective function values v.s. iterations numbers)