Numerical optimization: Assignment 10

Deadline: the lab on 2024.06.10 — THE FINAL LAB!

This time, we will deal with batch and stochastic gradient descents in a basic setup.

All the tasks will try to solve the following optimization problem.

Let there be $X_1, X_2, \ldots, X_n \in \mathbb{R}^d$ and $y_1, y_2, \ldots, y_n \in \mathbb{R}$ such that $F(X_i) = y_i$ for a certain **unknown** function F. Let there be a model function $G(\cdot; C)$ with parameters $C \in \mathbb{R}^m$. Find the parameters C which minimize a certain cost function.

You are not allowed to explicitly use neural networks.

1. $\boxed{\text{1 point}}$ Take the Iris dataset from the first two assignments and choose two types of irises. Define the optimization problem and the cost function if we want the function F to return 0 or 1, depending on the iris type. Propose a reasonable cost function. Propose a model function so that $m \leq n/3$.

2. $\boxed{\text{1.5 points}}$ Implement a batch gradient descent algorithm which uses all n observations in every step to calculate the full cost function for the full dataset. Use decay (very n steps, multiply the step size coefficient by $0 < \ell < 1$) and momentum, if necessary.

3. $\boxed{\text{1.5 points}}$ Implement a stochastic gradient descent algorithm which uses 10 random observations in every step to calculate the approximation of a cost function for the full dataset. Use decay and momentum, if necessary. Compare the iteration and convergence speeds.

4. $\boxed{\text{1 points}}$ Set aside at random 10% of the observations of each iris type (this will be a test dataset). Optimize the chosen model function on the remaining 90% of data. Then check how well the model function classifies the 10% of the irises which were set aside. The classification should assign the type 0 or 1 to each observation, depending on whether the model function's value is closer to 0 or to 1. Repeat the procedure 20 times, with new test sets and model optimization each time.