Projects

June 26, 2023

This document describes the project proposed for the lecture "Optimization in Machine Learning".

Stochastic Gradient Descent and its variants

The goal of this project is to experimentally compare the different variants of gradient descent algorithms in a simple classification setting.

Preparation

- You will need Git Large File Storage (Git LFS) for fetching the dataset from the Git repository. If you don't have it installed on your computer yet, you can do it following instructions there. The dataset is then downloaded by doing `git pull` in the `project/src` directory.
- It is a merged version of the `tabular-benchmark/covertype` dataset (through HuggingFaces). To load it, install the datasets library and PyTorch with `pip install datasets torch`. For instance, from the `src/` directory, load the downloaded dataset with

```
from datasets import load_from_disk

dset = load_from_disk("../data/covertype")
```

- Test the configuration with `python src/test.py`.

Training

With a one hidden-layer neural network (with the activation and width of your choice), compare (take at least one optional):

- Gradient descent
- (Optional) Stochastic gradient descent (SGD)

- Batch SGD
- (Optional) SGD + Momentum
- SGD + Nesterov Momentum
- (Optional) Adagrad
- Adam(W)
- (Optional) NAdam

If the training takes too much time, you can reduce the size of the dataset. Plot the results on a graph and comment them.