Neural Networks Hello World + Assignments 2, 3 (Neural Networks Implementation and Application Tutorial)

Vilém Zouhar, Noon Pokaratsiri Goldstein

24th, 25th November 2021

Overview

- Assignment 2
- Gradient
- PyTorch's Autograd
- NN Hello World
- Assignment 3

Assignment 2

- Tutor cue: go through the assignment
- What were the biggest issues? Coding or theory?
- Do you feel they are too easy/hard?
- Do you feel they are unrelated to the lecture content?

Optimization

Gradient 🤔



- What is it?
- How do we denote it?

$$abla f(p) = [rac{\delta f}{\delta imes_1}(p), \ldots, rac{\delta f}{\delta imes_k}(p)]$$

- Why is it important?
 - Optimalization

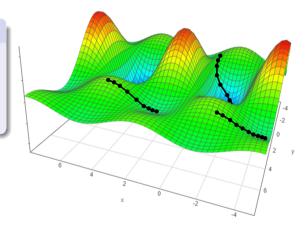


Figure 1: Function parameter landscape from [1]

Optimization

Few questions 🧐

- How does step/gradient-based optimization work?
- How is the step size determined?
- Why do we subtract the gradient and not add it?
- If we start in different places will we always find the same spot?
- Will we always find the global minimum?



Figure 2: Function parameter landscape from [2]

Autograd & PyTorch

How to get the gradient at (x, y) = (2, 3) of $x \cdot y + \sin(\pi \cdot x)$?

- By hand 😩
- Autograd

By hand

$$\frac{\delta}{\delta x} = y + \pi \cdot \cos(\pi \cdot x) \to 3 + \pi$$

$$\frac{\delta}{\delta y} = x \to 2$$

$$\nabla f(2,3) \to (3 + \pi, 2)$$

Autograd

```
import torch
import numpy as np
x = torch.tensor(2.0, requires_grad=True)
y = torch.tensor(3.0, requires_grad=True)
out = x*y + torch.sin(np.pi*x)
out.backward() # trigger gradient computation
assert np.isclose(x.grad, 3+np.pi)
assert np.isclose(y.grad, 2)
```

Assignment 3

• Any questions?

Resources

- $\bullet \ [1] \ Optimization \ \& \ landscapes \ offconvex.org/2018/11/07/optimization-beyond-landscape/$
- [2] Optimization Introduction by Scipy scipy-lectures.org/advanced/mathematical_optimization/