COMP9444 Neural Networks and Deep Learning Term 2, 2020

Solutions to Exercises 4: PyTorch

This page was last updated: 06/22/2020 14:15:01

Download the zip file `Ex4code.zip` and unzip it. This will create a directory `Ex4code` with two very simple PyTorch programs, `slope.py` and `xor.py`.

---

1. **Adjusting the Learning Rate and Momentum**

   The program `slope.py` solves the simplest possible machine learning task:

   solve `F(x) = A*x` such that `F(1)=1`

   a. Run the program by typing

   ```
   python3 slope.py --lr 0.1
   ```

   Try running the code using each of the following values for learning rate:

   `0.01, 0.1, 0.5, 1.0, 1.5, 1.9, 2.0, 2.1`

   Describe what happens in each case, in terms of the success and speed of the algorithm.

   ```
   0.01   the task is learned in 998 epochs
   0.1    the task is learned in 97 epochs
   0.5    the task is learned in 16 epochs
   1.0    the task is learned in   2 epoch
   1.5    the task is learned in  16 epochs
   1.9    the task is learned in  97 epochs
   2.0    the parameter oscillates between 0.0 and 2.0
   2.1    the parameter explodes, first to inf and then to nan
   ```

   b. Now add momentum, by typing:

   ```
   python3 slope.py --mom 0.1
   ```

   Try running the code with learning rate kept at the default value of `1.9` but with momentum equal to `0.1, 0.2,` etc. up to `0.9`. For which value of momentum is the task solved in the fewest epochs?

   > The fewest is `13` epochs, when momentum is `0.3`.

   What happens when the momentum is `1.0`? What happens when it is `1.1`?

   > When momentum is `1.0`, the Parameter cycles around and never converges. When momentum is `1.1`, the Parameter blows up to Infinity.

2. **Learning the XOR Task**

   The program `xor.py` trains a 2-layer neural network on the XOR task.

a. Run the code ten times by typing

```
python3 xor.py
```

For how many runs does it reach the global minimum? For how many runs does it reach a local minimum?

It reaches the global minimum in approximately 50% of runs; it gets stuck in a local minimum for the remaining 50%.

b. Keeping the learning rate fixed at `0.1`, can you find values of momentum (`--mom`) and initial weight size (`--init`) for which the code converges relatively quickly to the global minimum on virtually every run?

We have found that these hyperparameters successfully reach the global minimum in 99% of runs:

```
python3 xor.py --mom 0.9 --init 0.01
```
