

6장. 학습 관련 기술들(1) 실습문제

신경망

학번: 202404178 이름: 강우현

※ hwp 파일 한 개, py 파일 3개 제출하세요.

1. 6.1 절에서 소개한 매개변수 업데이트 방법 4가지의 이름, 수식, 특징을 간단히 적으시오.

확률적 경사 하강법(SGD) : (수식) $W \leftarrow W - \eta \frac{\partial L}{\partial W}$ / (특징) 단순하고 구현이 쉽지만, 문제에 따라 비효율적일 때가 있다.

모멘텀(Momentum) : (수식) $v \leftarrow \alpha v - \eta \frac{\partial L}{\partial W}$, $W \leftarrow W + v$ / (특징) 어떤 물체가 움직일 때 계속 그 방향으로 움직이게 되는 경향을 모멘텀이라 한다.

AdaGrad(Adaptive Gradient) : (수식) $h \leftarrow h + \frac{\partial L}{\partial W} \odot \frac{\partial L}{\partial W}$, $W \leftarrow W - \eta \frac{1}{\sqrt{h}} \frac{\partial L}{\partial W}$ / (특징)

최솟값을 향해 효율적으로 움직인다. 지그재그 움직임이 많이 줄어들

Adam(Adaptive Moment) : (특징) 모멘텀과 AdaGrad 기법을 합친 것.

※ C 드라이브 아래 Neural_Net 폴더 아래, ch06 폴더를 만들고 아래 문제 2번, 3번 코드를 작성하시오. 교재 Github(<https://github.com/WegraLee/deep-learning-from-scratch/tree/master>)에서 필요한 파일을 다운로드하여 적절한 폴더에 넣으시오.

2. 6.1.7 절의 optimizer_compare_naive.py 코드를 다음과 같이 수정하여 실행하시오.

- 목적 함수를 $f(x,y) = \frac{1}{10}x^2 + y^2$ 로 수정하시오.
- 초깃값(initial value)을 (-5, 1)로 변경하시오.
- 50번 반복하시오.
- optimizer마다 반복 후, optimizer 이름과 그 x_history와 y_history 값을 적절히 출력하시오.

3. 6.1.8 절의 optimizer_compare_mnist.py 코드를, 100개 노드를 가지는 은닉층을 3개를 두도록 수정하여 Visual Studio에서 실행하시오. 어느 optimizer가 좋은지 판단하시오. 왜 좋은가?

AdaGrad가 학습이 제일 빠르다.

4. (팀플 관련) Neural_Net 폴더 아래 data 폴더를 만들고 작성하시오.

(1) <https://github.com/zalandoresearch/fashion-mnist/tree/master> 에서 data/fashion/ 폴더에서 4개의 압축 파일을 다운받으시오.

(2) Neural_Net/data/fashion/ 에 (1)에서 다운로드한 4개 파일을 넣으시오.

(3) 위 사이트에서 mnist_reader.py 파일을 다운로드하고 data 폴더 안에 넣으시오

(4) optimizer_compare_mnist.py 파일을 fashion mnist 데이터에 대해 실행하는 것으로 코드를 수정하여 optimizer_compare_fashion_mnist.py에 저장하여 제출하시오.

(힌트)

```
from data import mnist_reader  
x_train, t_train = mnist_reader.load_mnist('../data/fashion/', kind='train')  
x_test, t_test = mnist_reader.load_mnist('../data/fashion/', kind='t10k')
```

팀별로 상의하여 답안을 작성하세요.