

from-scratch-optimization

☰ Tags

From scratch Chap6. 학습 관련 기술들

- 밑바닥부터 시작하는 딥러닝 - 사이토 고키지
- 신경망 학습의 핵심 개념 재정리

Agenda

- 가중치 매개변수 최적값 탐색 최적화 방법
- 가중치 매개변수 초기값
- 하이퍼파라미터 튜닝
- weight decay
- dropout
- batch normalization

6.1. 매개변수 갱신

Q. 신경망 학습의 목적은 무엇일까?

- A. 신경망 학습의 목적은 손실 함수의 값을 가능한 한 낮추는 매개변수를 찾는 것이다
- 즉, 매개변수의 최적값을 찾는 문제이다. 이를 최적화(Optimization)라고 한다

Q. 지금까지 배운 최적화 방법은 어떤 방법이 있을까?

- A. 지금까지는 최적의 매개변수를 찾는 단서로 매개변수의 기울기(미분)을 이용했다
- 매개변수에 대한 손실함수의 기울기를 구해, 기울어진 방향으로 매개변수 값을 갱신하는 일을 계속해서 반복하여 최적의 값에 다가갔다
- 이는 확률적 경사 하강법(SGD)라고 한다

- SGD

- $W \leftarrow w - \alpha * dl/dw$

```
class SGD:
    def __init__(self, lr=0.01):
        self.lr = lr

    def update(self, params, grads):
        for key in params.keys():
            params[key] -= self.lr * grads[key]
```

- pseudo code

```
network = TwoLayerNet(...)
optimizer = SGD()

for i in range(10000):
    ...
    x_batch, t_batch = get_mini_batch(...). # 미니배치
    grads = network(x_batch, t_batch)
    params = network.params
    optimizer.update(params, grads)
    ...
```