

2023_11_08 DSL 스터디

4. 신경망 학습

1. 손실함수

1) 오차제곱합

평균 제곱 오차(mean squared error, MSE)라고도 한다.

$$E = \frac{1}{2} \sum_k (y_k - t_k)^2$$

- y_k 신경망의 출력(신경망이 추정한 값)
- t_k 정답 레이블
- k 데이터의 차원 수

실제 정답과 예측값 사이의 차이를 수치화 한 것이므로, 값이 작을수록 좋은 지표

```
def mean_squared_error(y, t):  
    return 0.5 * np.sum((y-t)**2)  
  
# 정답 레이블(실제 정답은 2)  
t = [0, 0, 1, 0, 0, 0, 0, 0, 0, 0]  
  
# 예1: '2'일 확률이 가장 높다고 추정함 (0.6)  
y = [0.1, 0.05, 0.6, 0.0, 0.05, 0.1, 0.0, 0.1, 0.0, 0.0]  
mean_squared_error(np.array(y), np.array(t))  
>>>  
0.09750000000000003  
  
# 예2 '7'일 확률이 가장 높다고 추정함 (0.6)  
y = [0.1, 0.05, 0.1, 0.0, 0.05, 0.1, 0.0, 0.6, 0.0, 0.0]  
mean_squared_error(np.array(y), np.array(t))  
>>>  
0.5975
```

- 정답과 멀어질수록 MSE가 커진다.

밑바닥부터 시작하는 딥러닝 4단원을 학습했다.

GA 동작 원리

동작 원리를 간단하게 설명하면,

먼저, random한 초기 염색체(최적의 해가 될 후보들)를 생성하고, 이 중 **자연 선택**된 염색체들끼리 **교차(Crossover)** & **변이(Mutation)** 연산을 하여 자식 세대를 생성한다. 이 과정을 반복하면서 점차 최적의 해를 찾아간다.

유전알고리즘이 어떻게 동작하는지 자세한 과정은 각 step 별로 R로 구현한 코드와 함께 설명하겠다.

R코드는 [이 사이트](#)를 참고하였다. (참고한 사이트에서는 코드에서의 값과 설명을 위한 예시가 달라 좀 헷갈렸다. 본 포스팅에서는 코드의 값과 설명을 위한 예시를 같이 하여 진행하겠다.)

Step 0. 문제 정의

먼저, 최적화 하고 싶은 문제를 정의한다.

문제는 다음과 같다.

$2a^2 + b = 57$ 를 만족하는 최적의 a와 b를 추정해보자.

국소 최솟값을 찾을 수 있는 경사 하강법의 한계에 대해 알아보고 global minimum을 찾을 수 있는 유전 알고리즘에 대해 학습하였다.