

Pseudo Gradient Descent

April 27, 2022

비문 다듬기

1 문제

산지니가 수강하고 있는 머신러닝 강의에서 Gradient Descent를 수행하는 과제가 출제되었다. Gradient Descent란 머신러닝 모델의 오차를 나타내는 Loss Function이 최소가 되도록 모델의 매개 변수를 학습하는 것을 지칭하였고 이것을 2개의 매개 변수를 가진 머신 러닝 모델의 Loss Function을 시각화한 아래 그림으로 설명하자면 Gradient Descent에서 매개 변수의 학습은 Loss Func-

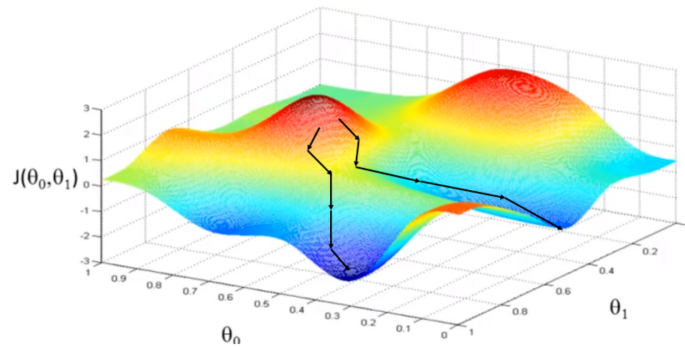


Figure 1: Gradient Descent

tion을 각 매개변수에 대하여 편미분을 수행하여 Loss Function의 기울기가 낮아지는 방향으로 학습하는 것을 말한다. 하지만 순수한 Gradient Descent는 항상 기울기를 따라 학습이 이루어지기에 다음과 같은 Local Optimum에 빠지면 탈출할 수 가 없다.

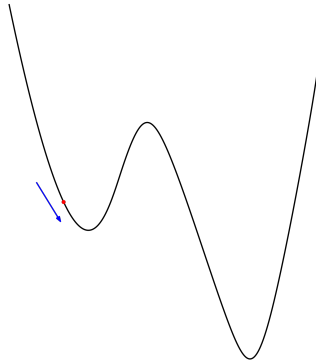


Figure 2: Local Optimum

그래서 Gradient Descent에 momentum이라는 일종의 관성력을 추가해서 사용하기도 하는데, momentum은 마치 현실의 관성력이 그러하듯 과거에 진행해왔던 방향으로 계속 진행하려는 것을 말하며 이 덕분에 일시적인 머신러닝 모델의 학습과정에서 global optimum에 빠지더라도 모델은 어느 정도 탈출할 수 있게 된다.

이때 산지니가 해야할 것은 2개의 매개 변수를 가지는 머신러닝 모델에 momentum을 적용한 gradient descent를 적용하여 모델의 학습 경로를 기록하고 제출하는 것이다. 하지만 산지니는 미분을 할 줄 모르기에 Loss Function의 공간을 전역탐색하고 그걸 바탕으로 그럴듯 하게 구현하기로 했다.

산지니는 경로를 구할때 Gradient Descent의 정의대로 현재 위치에서 Loss Function의 값이 최소가 되는 방향으로 매개변수 하나를 1만큼 변화하도록 하기로 했으며 momentum의 경우 임의로 이전 진행방향대로 매개변수를 1만큼 변화하도록 하였다. 이때 Gradient Descent로 인한 매개변수 업데이트와 momentum으로 인한 매개변수 업데이트는 동시에 일어날 수 없도록 하였으며 또한 연속해서 momentum을 적용하지 않기로 했다.

하지만 산지니는 전역변수의 공간을 탐색하는 것에 너무 힘을 쓴 나머지 그만 잠이 들어 함수 공간은 전부 찾았으나 경로 탐색은 수행하지 못하였다. 지금 부터 여러분이 수행할 일은 산지니를 대신하여 경로를 찾아주는 것이다.

2 입력

먼저 산지니가 찾아놓은 Loss Function 공간의 크기 N , M 이 주어진다. 그리고 다음 N 개의 줄에 걸쳐 M 개의 Loss Function의 값이 주어진다. 마지막으로 는 경로 탐색을 시작할 위치 ω_0 , ω_1 가 주어진다. 이때 ω_0 는 공간의 x 축, ω_1 은 공간의 Y 축을 의미하며 좌측 최상단은 각각 0 일때를 의미한다.

3 출력

경로에 따른 ω_0 , ω_1 의 값 한 줄 씩 출력하면 된다.

4 예제 입력

6 7

100	101	200	150	300	400	1
87	89	65	67	66	67	100
30	38	21	40	55	57	300
40	35	18	20	10	10	10
43	41	6	4	10	2	4
100	100	100	100	10	4	4

0 0

5 예제 출력

0 0

0 1

0 2

0 3

1 3

2 3

2 4

3 4

4 4

5 4

6 알고리즘

Greedy, DP, BFS

7 예외 케이스

처음부터 못 움직이는 경우

momentum이 공간 밖으로 탈출하는 경우