

Adam Optimizer

practice 인공지능,머신러닝 · 2024. 10. 12. 11:54

Q.딥러닝에서 Gradient Descent의 변형인 Adam Optimizer의 특징과 장점에 대해 설명해 주세요.

Adam optimizer는 지속적으로 학습률을 조절해, 더 빠르고 안정적인 학습을 가능하게 합니다. 처음에는 optimum (맨 밑바닥)에 멀리 떨어져 있을 가능성이 높으므로, 높은 학습률을 적용합니다.

그리고 학습이 진행될수록 (optimum에 가까워 질수록) 학습률을 낮춥니다 (optimum을 지나치지 않도록)

여기에 추가적으로, Adam은 모멘텀 (관성)을 사용합니다.

만약 같은 방향으로 계속해서 비슷한 속도로 이동하고 있다면, optimum과 멀리 떨어져 있는 것입니다.

이러한 상황에서는 학습률을 감소시키지 않거나 약간 증가시킵니다.

반대로 이동 방향이 자주 바뀌고, 기울기가 작다면 optimum 근처까지 온 것이므로 학습률을 감소시킵니다.

•딥러닝에서 Gradient Descent의 변형인 Adam Optimizer의 특징과 장점에 대해 설명해 주세요.

•역전파를 통해 Global Optimum을 찾는 것이 목표

•처음 시작점은 optimum(맨밑바닥)과 비교했을 때, 멀리 떨어져 있을 가능성이 높음
상대적으로 높은 학습률로 시작

- 이후 학습이 진행되면서, 점차 optimum에 가까워 질 것임
점진적으로 학습률을 낮춰, optimum을 지나치지 않도록 함

- 이때 Adam은 추가적으로 Momentum(관성)을 사용함

- 만약 같은 방향으로 계속 이동한다면, 학습률을 높게 유지

- 만약 학습 기울기가 전반적으로 비슷하고, 상대적으로 크다면 지금 optimum과 매우 멀리 떨어져 있는 것임

보다 높은 학습률

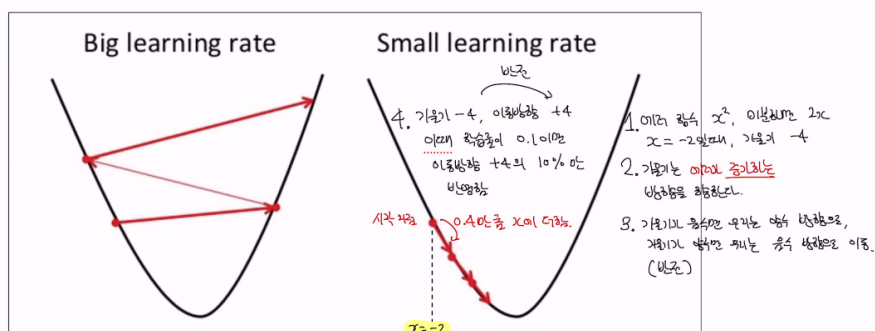
- 반대로, 학습 기울기가 계속해서 많이 바뀌고, 크기가 작아졌다면 지금 optimum 근처까지 온 것임

학습률을 점차 낮춤

아래그림) 에러함수

Learning Rate

- 만약 $\text{Error} = x^2$ 일때



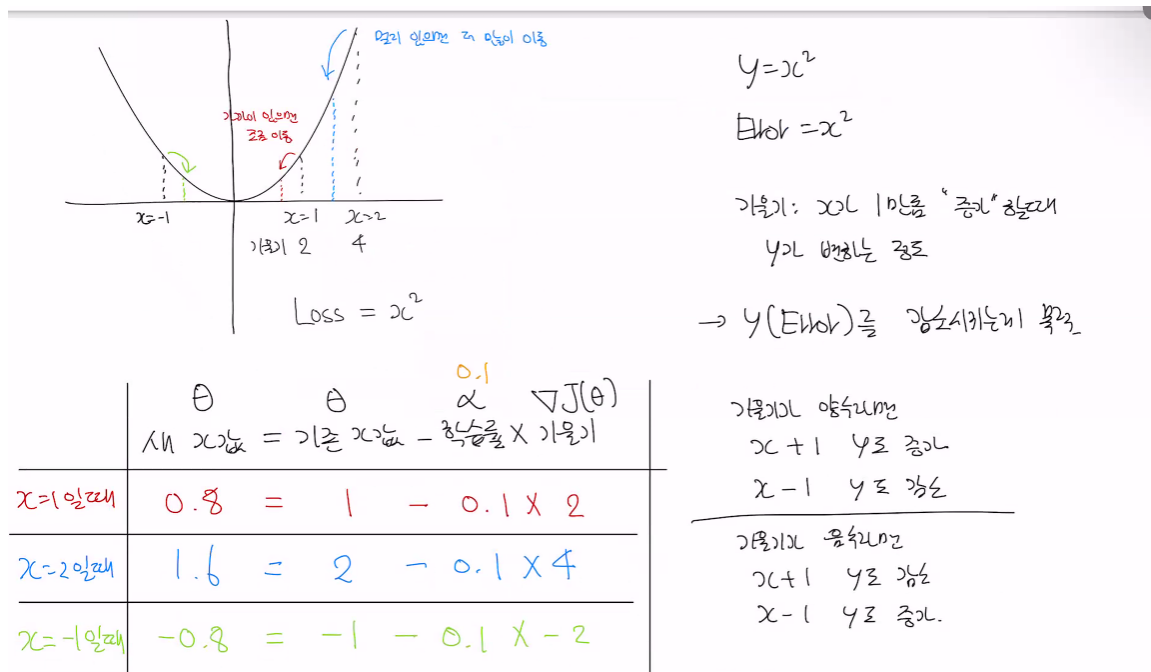
출처: 딥러닝·인공지능 Signature 초격차 패키지 Online.

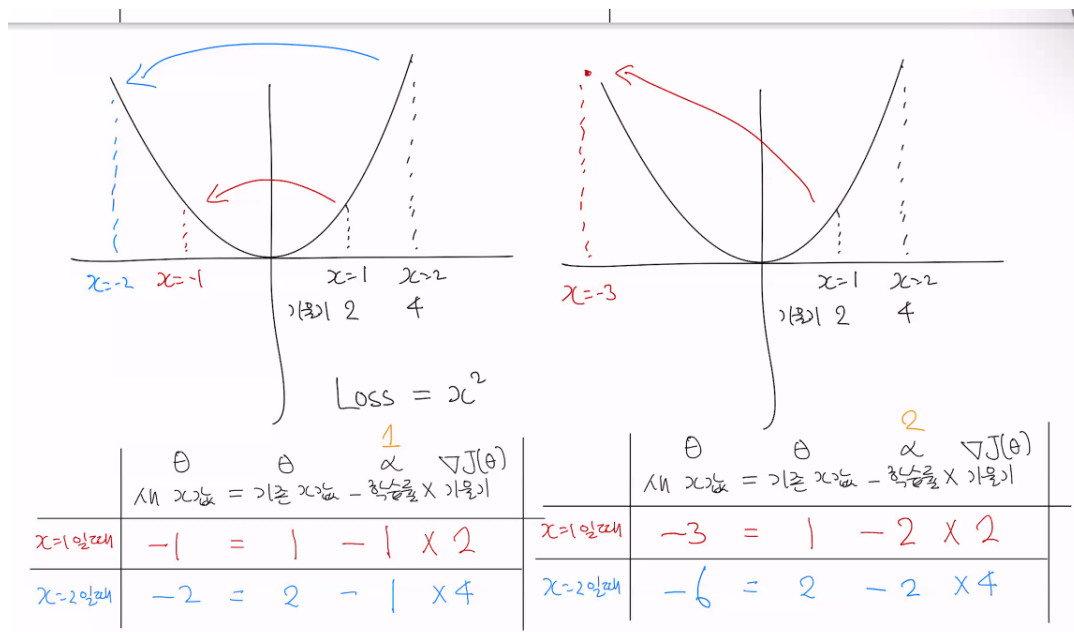
경사하강법

경사하강법(Gradient Descent)은 머신러닝과 최적화 문제에서 자주 사용되는 알고리즘 중 하나로, 목표는 비용 함수(또는 손실 함수)를 최소화하는 변수들의 최적 값을 찾는 것입니다. 경사하강법은 반복적으로 비용 함수의 기울기를 계산하고, 그 기울기의 반대 방향으로 이동함으로써 점차적으로 최적 해를 찾아가는 방식입니다.

경사하강법의 기본 개념

1. ****목표****: 비용 함수를 최소화하는 파라미터(예: 가중치, 편향 등)를 찾는 것.
2. ****기울기****: 특정 점에서 비용 함수의 기울기를 계산해, 그 방향이 비용을 증가시키는지 감소시키는지 확인.
3. ****업데이트****: 기울기의 반대 방향으로 파라미터를 업데이트함으로써 점점 최소값에 가까워짐.





♡ 2



구독하기

'practice 인공지능,머신러닝' 카테고리의 다른 글

GPT3 (0)

2024.10.12

Transformer, RNN 차이 (0)

2024.10.12

Batch Normalization (3)

2024.10.12

강화학습 (0)

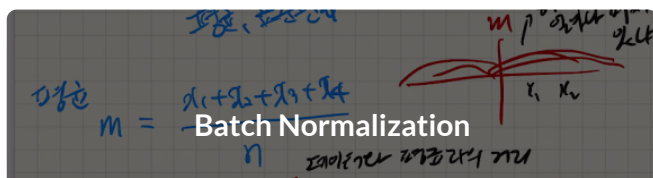
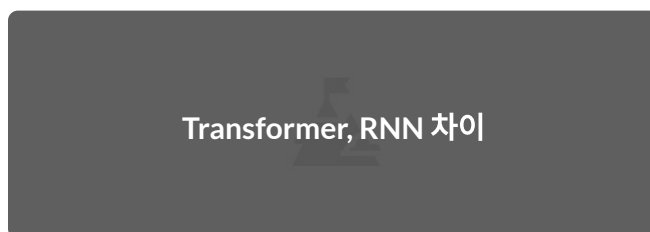
2024.10.12

앙상블 학습 (0)

2024.10.11

관련글

관련글 더보기



4. 주어진 식 = $\sqrt{(x_1-m)^2 + (x_2-m)^2 + (x_3-m)^2 + (x_4-m)^2}$

자연어(NLP)

네이쳐2024 님의 블로그입니다.



구독하기 +

댓글 1



익명



비밀댓글입니다.

2024. 10. 12. 12:13



이름

비밀번호

내용을 입력하세요.



이름