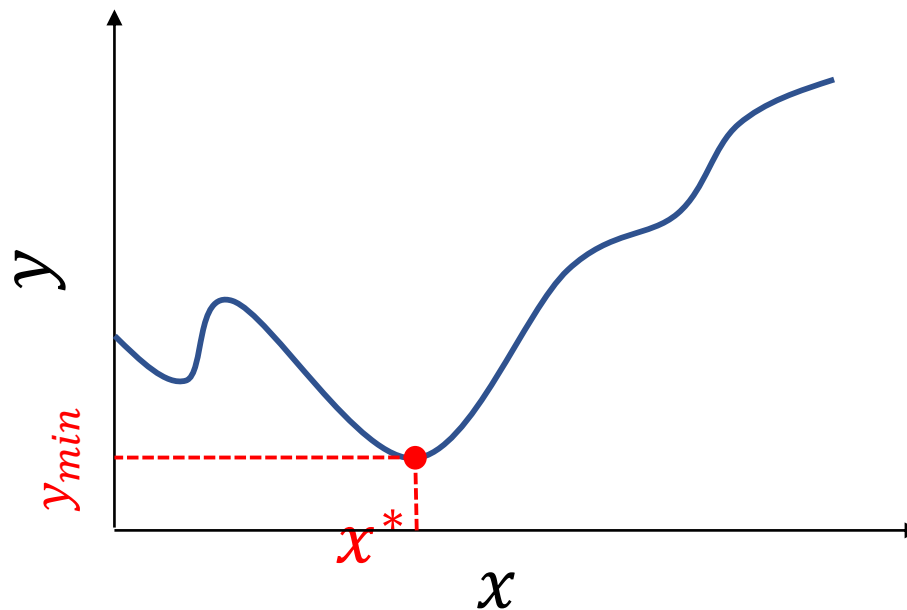


Chapter 03. 쉽게 배우는 경사 하강 학습법

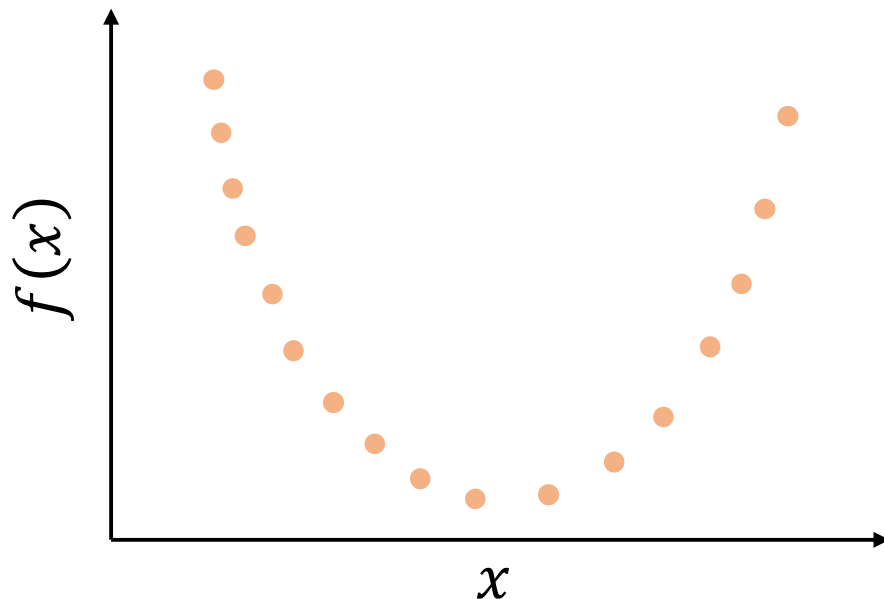
STEP1. 경사 하강 학습법

최적화 이론



손실 함수를 최소화 하는 입력 값(최적 값)을 찾아내는 연구
전쟁 물자의 분배를 연구하면서 발전

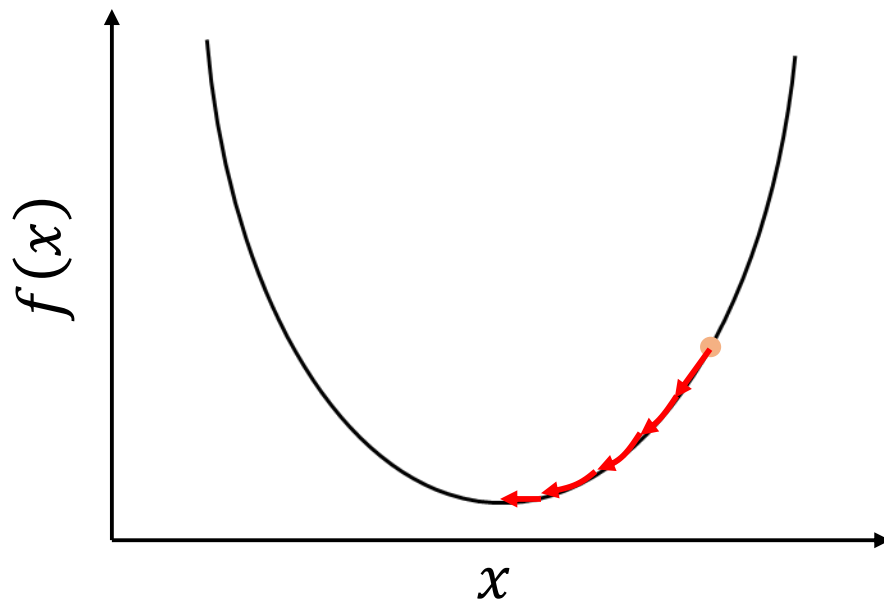
무차별 대입법 Brute-Force



- 가능한 모든 수를 대입해 보는 방법
- 가장 단순한 방법으로 함수를 알 수 있음
- 다음과 같은 문제로 최적화에 이용할 수 없다.
 - x^* (최적값)이 존재하는 범위를 알아야 함
 - x^* 를 정확히 찾기 위해 무한히 촘촘하게 조사해야 함
 - $f(x)$ 의 계산 복잡도가 매우 높음

적게 대입해 보고 답을 찾을 수는 없을까?

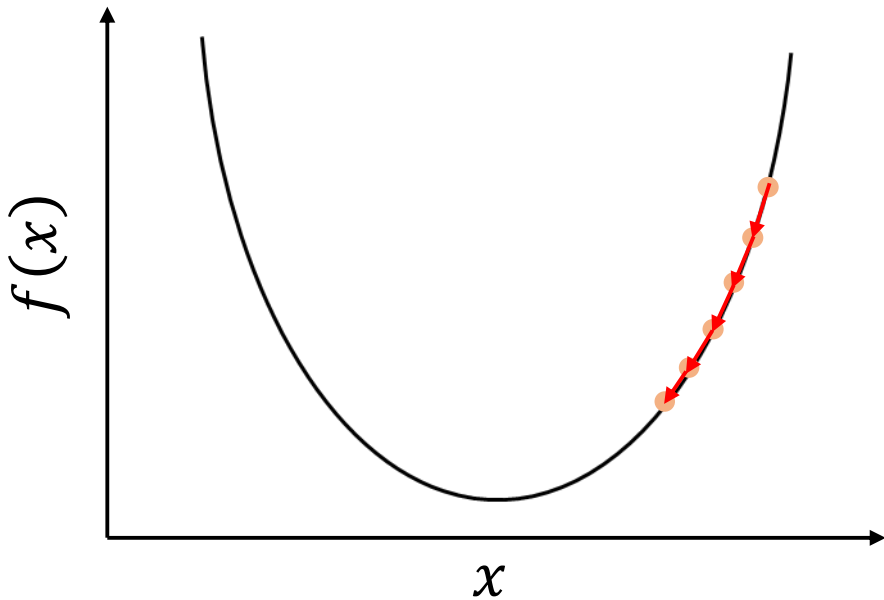
경사 하강법 Gradient Descent



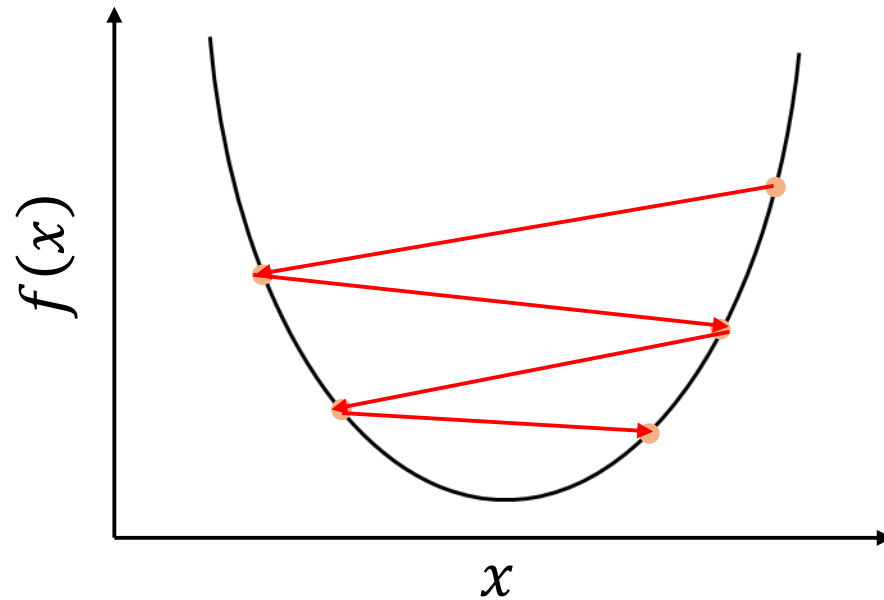
경사를 따라 여러 번의 스텝을 통해 최적점으로 다가간다.

경사는 기울기(미분, Gradient)를 이용해 계산한다.

학습률의 선택



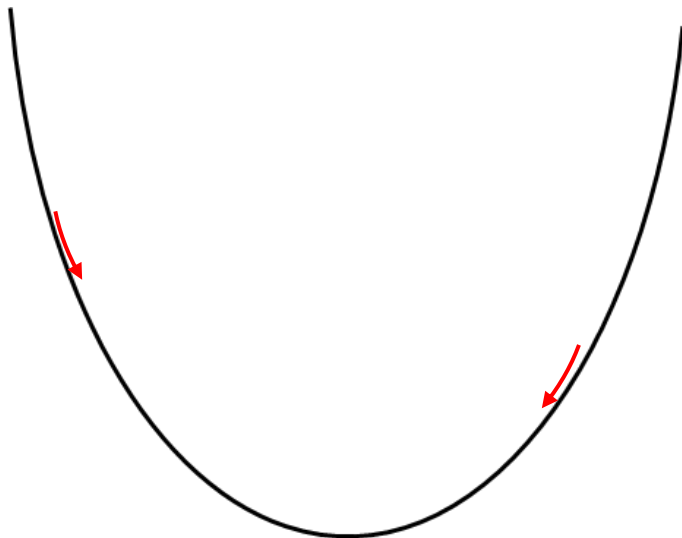
α 가 너무 작은 경우



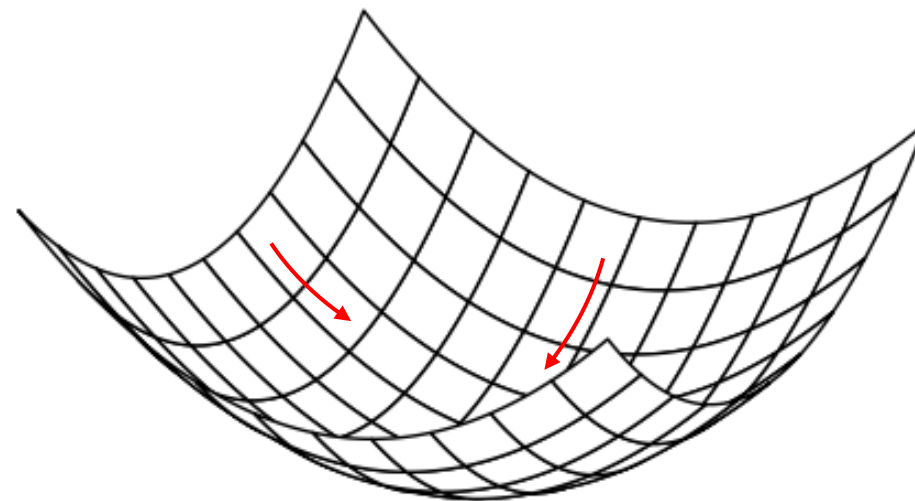
α 가 너무 큰 경우

학습률(Learning Rate) α 에 비례하여 이동한다. 적절한 학습률을 선택하는 것은 매우 중요하다.

볼록 함수 Convex Function



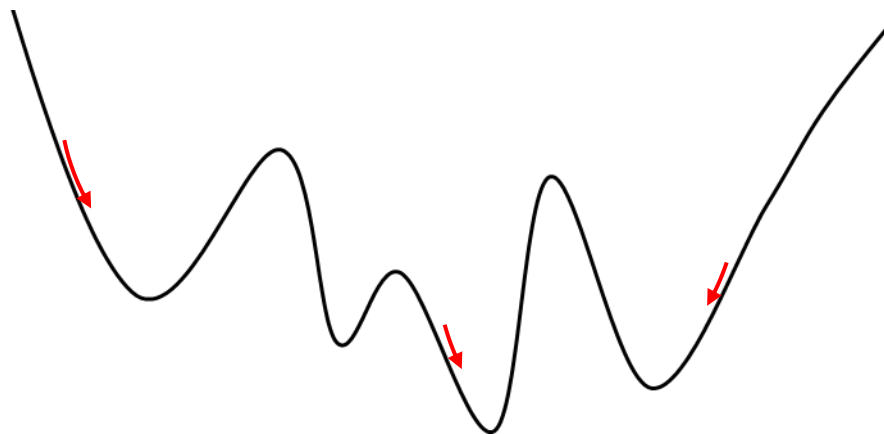
1-D Convex function



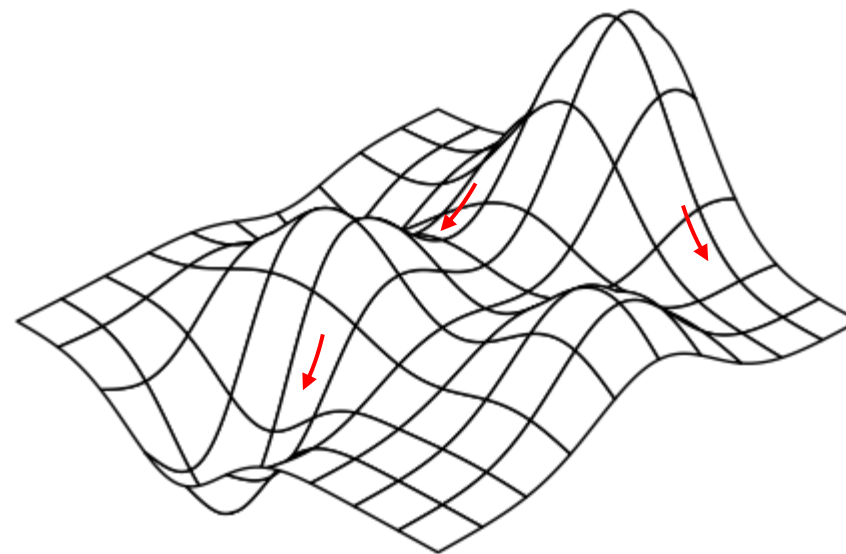
2-D Convex function

볼록 함수(Convex function)는 어디서 시작하더라도 경사 하강법으로 최적 값에 도달할 수 있다.

비볼록 함수 Non-convex Function



1-D Non-convex function



2-D Non-convex function

비볼록 함수(Non-convex function)는 시작 위치에 따라 다른 최적 값을 찾는다.
즉, 지역 최적값(Local Minimum)에 빠질 위험이 있다.