

## 4.4 | 경사법(경사 하강법)

1. 학습 시 최적의 매개 변수(손실함수가 최소값이 될 때의 매개 변수)를 찾아야 함

2. 하지만 일반적으로 매개 변수 공간이 광대하여  
어디가 최소값이 되는 곳인지 짐작할 수 없음

3. 이런 상황에서 기울기를 잘 이용해 함수의 최소값  
(또는 가능한 작은 값)을 찾으려는 것이 '경사법'임

4. 하지만 기울기가 가리키는 방향으로 최소값이 있는지 대략은

5. #1 방법으로 '최소값'을 가리키는 것은 아니나, 그 방향으로 가야  
함수의 값을 줄일 수 있음

6. 그러므로 기울기 정보를 단서로 나아갈 방향을 짐작해야 함

7. 이렇게 함수의 값을 줄이는 것을 **경사법**이라 함.

\* 속식

$$\theta_0 = \theta_0 - \eta \frac{\partial f}{\partial \theta_0}, \quad \theta_1 = \theta_1 - \eta \frac{\partial f}{\partial \theta_1}$$

$\eta^{\text{eta}}$  : 학습률 (갱신하는 양)

→ 너무 크거나 작으면 '좋은 장소'를 찾아갈 수 없음

보통 이 학습률 값을 변경하며 올바르게 학습중인지 확인함

\* 학습률이 너무 크면 큰 값으로 발산,

너무 작으면 갱신되지 않은 채 끝남.