

Math for NN

Gradient Descent Optimization

- Error(오차 제곱의 합)을 최소화 하기 위한 방법

Solve

Find w_1 which minimizes the following error function

$$E(w_1) = w_1^2 + 2w_1 + 3$$

- Too easy

1. Differentiate E with respect to w_1

$$\frac{dE(w_1)}{dw_1} = 2w_1 + 2$$

2. If you find w_1 that makes this 0

$$2w_1 + 2 = 0$$

3. That's the solution!!

$$w_1 = -1$$

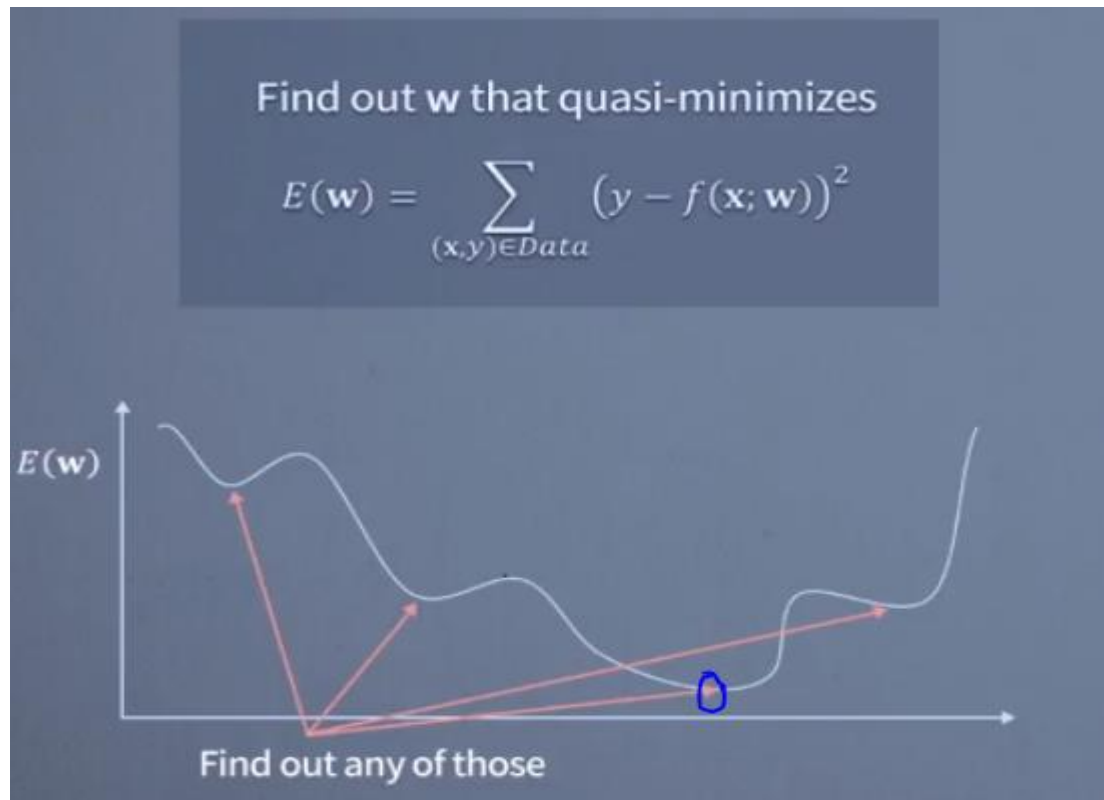
W가 1개인 경우 w에 대해 미분한 뒤, 미분한 식 =0을 통해 w를 찾아낸다.

W가 여러 개인 경우 각각 w에 대해 편미분한 뒤, 연립방정식을 통해 동시에 0이 되는 w를 찾아낸다.

1차식 함수에 대해 풀기 쉽지만, 다른 함수에 대해서 해를 찾기 굉장히 어려운 문제이다.

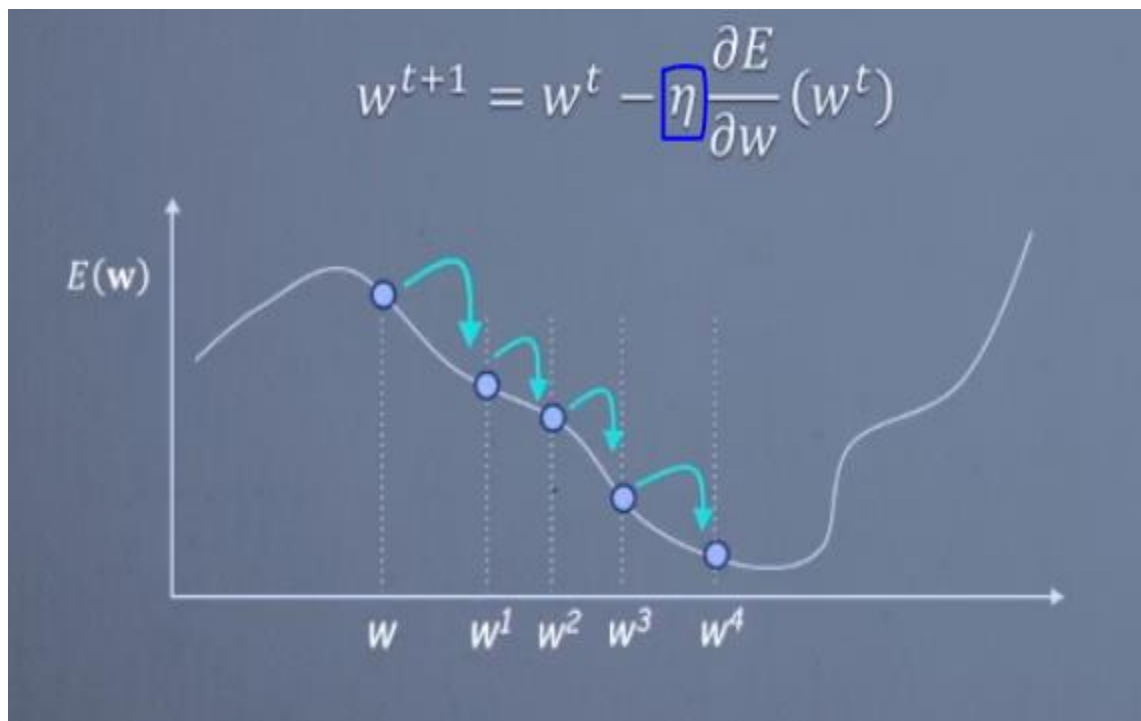
Quasi-minimize

- Error(오차 제곱의 합)을 최소화 하기 위한 방법



최소화하는 지점 1개만 찾는 것이 아니라 여러 개 존재하는 local minimum을 찾는다.

Gradient Descent Optimization



1. 임의의 지점을 Random하게 잡는다.
 2. 기울기를 계산한다.
 3. 최소값에 도달할 때 까지 계속 지점을 이동한다.(기울기 + : minus 방향으로, 기울기 - : positive 방향으로)
 4. 기울기가 0에 도달할 때 까지 계속 반복한다.
- Learning rate : 직접 주는 상수 값
 - w 값이 여러 개라면 각 w_0, w_1, w_2 마다 편미분을 하고, 위 절차를 반복한다.

편미분

$$\begin{array}{l} y = f(x) \\ x = g(t) \end{array} \quad \frac{dy}{dt} = ? \quad \frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt}$$

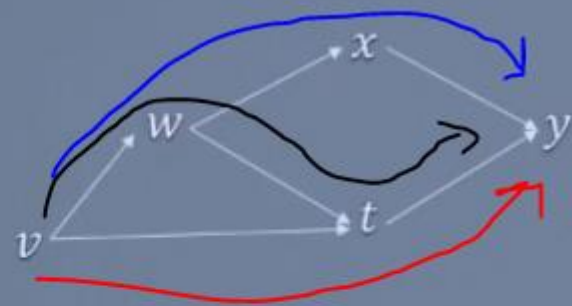
$$t \longrightarrow x \longrightarrow y$$

$$\begin{array}{l} y = f(x) \\ x = g(t) \\ t = h(w) \end{array} \quad \frac{dy}{dw} = ? \quad \frac{dy}{dw} = \frac{dy}{dx} \frac{dx}{dt} \frac{dt}{dw}$$

$$w \longrightarrow t \longrightarrow x \longrightarrow y$$

$$\begin{array}{l} y = f(x, t) \\ x = g(w) \\ t = h(w, v) \\ w = i(v) \end{array}$$

$$\frac{\partial y}{\partial v} = ?$$



$$\frac{\partial y}{\partial v} = \frac{\partial y}{\partial x} \frac{dx}{dw} \frac{dw}{dv} + \frac{\partial y}{\partial t} \frac{\partial t}{\partial w} \frac{dw}{dv} + \frac{\partial y}{\partial t} \frac{\partial t}{\partial v}$$