

AI and Deep Learning

4. Linear Regression & Influence on error

Yung-Cheol BYUN, Jeju National Univ.

여행, 그리고 회귀

회귀(Regression)

인류는 고향을 떠나도 나이가 들면
언젠가는 본래의 고향으로
회귀하고(돌아가고) 싶어한다.
(인류학)



자연의 법칙, 섭리, 일종의 규칙

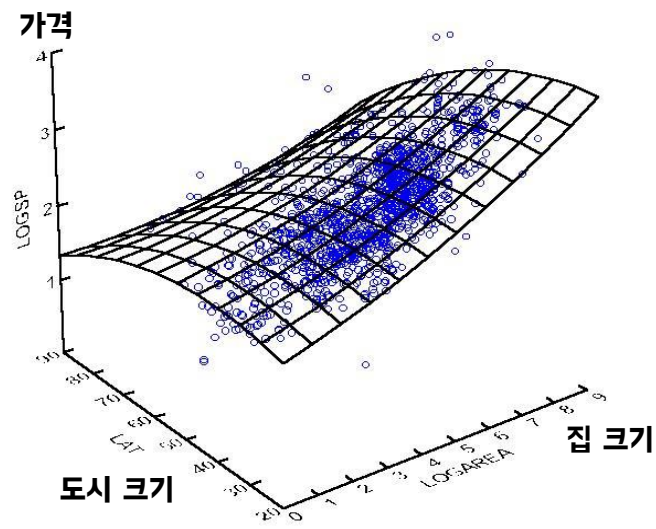
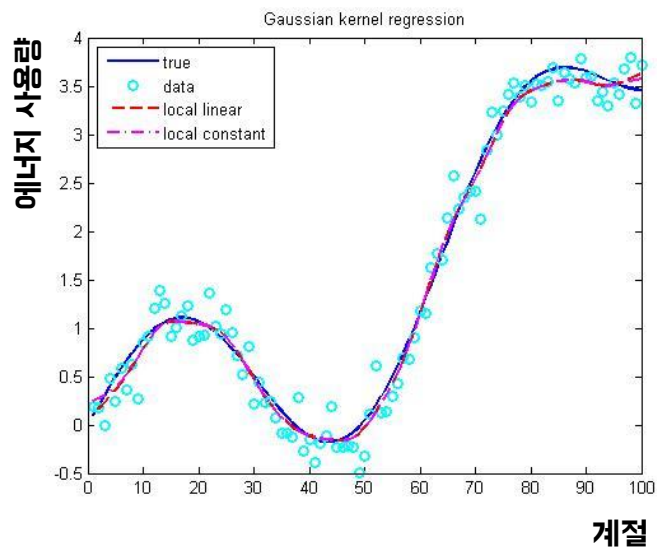
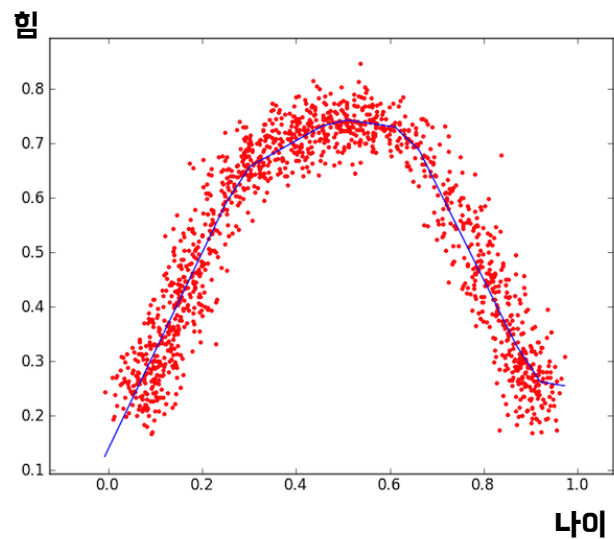
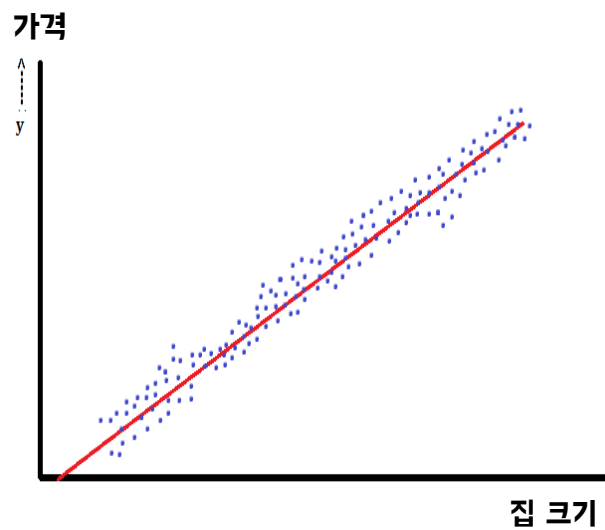
- 연어는 태어난 곳으로 돌아온다.
- 집은 클수록 비싸다.
- 젊을 때는 강하지만 어리거나 늙으면 약하다.
- 남자가 여자보다는 큰 편이다.
- 성적이 좋을 수록 취업이 잘된다.

반드시 그런 것은 아니지만 일반적으로 그런 경향이 있다.

이를 잘 표현하는 말, 용어

회귀(Regression)

‘집이 클수록 가격이 비싸다.’ (회귀)
집의 크기와 가격 간의 관계
회귀는 그래프로 표현하면 이해하기 쉬움.



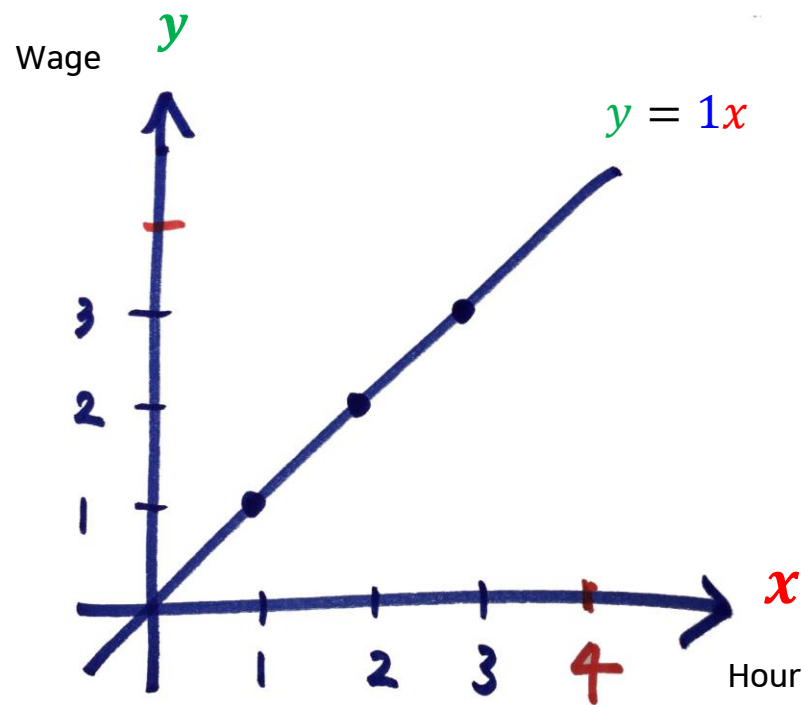
회귀(Regression)

하나 이상의 독립변수(X)와
하나의 종속변수(Y) 사이의 관계
(statistical measure to determine
the relationship)

Linear Regression

The relationship forms **linear** shape.
ex) wage/hour, price/size of house

가격



www.desmos.com

Goal

- 데이터는 Linear Regression
경향
- 모든 점을 지나는 직선 찾기
- 혹은 가장 근사하게 지나는
직선 찾기

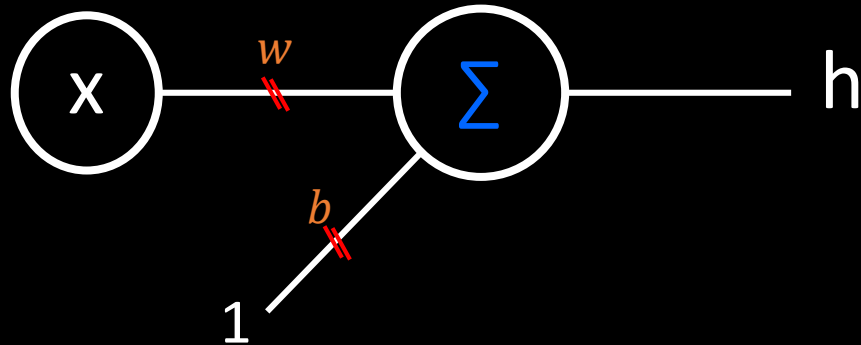
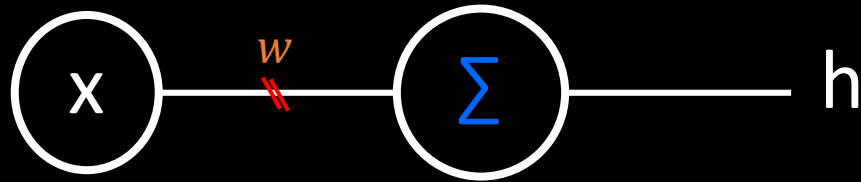
가설(Hypothesis)

$$h = wx$$

$$h = wx + b$$

- 가설은 아직 증명되지 않으나 조정을 통하여 데이터를 표현할 수 있다고 믿는 것 (가정)
- 반복적으로 w (and b) 값을 적절히 조정할 경우 우리의 데이터를 잘 표현할 수 있을 것으로 믿음.

Bias를 갖는 뉴런



어떻게 w 를 조장할 것인가?

- 모든 점을 지날 경우 차이 값(오류 , 에러, 비용, loss)은 0
- 따라서 차이 값이 0이 되도록 w 와 b 를 조금씩 계속해서 조정

오류(e) = |뉴런이 예상한 값 - 정답|

$$e = |wx - y|$$

데이터가 1개($x=1, y=1$)라면?

$$e = |w \cdot 1 - 1|$$

자, 이제 e 가 0이 되는 w 를 찾자.

x	y
1	1
2	2
3	3

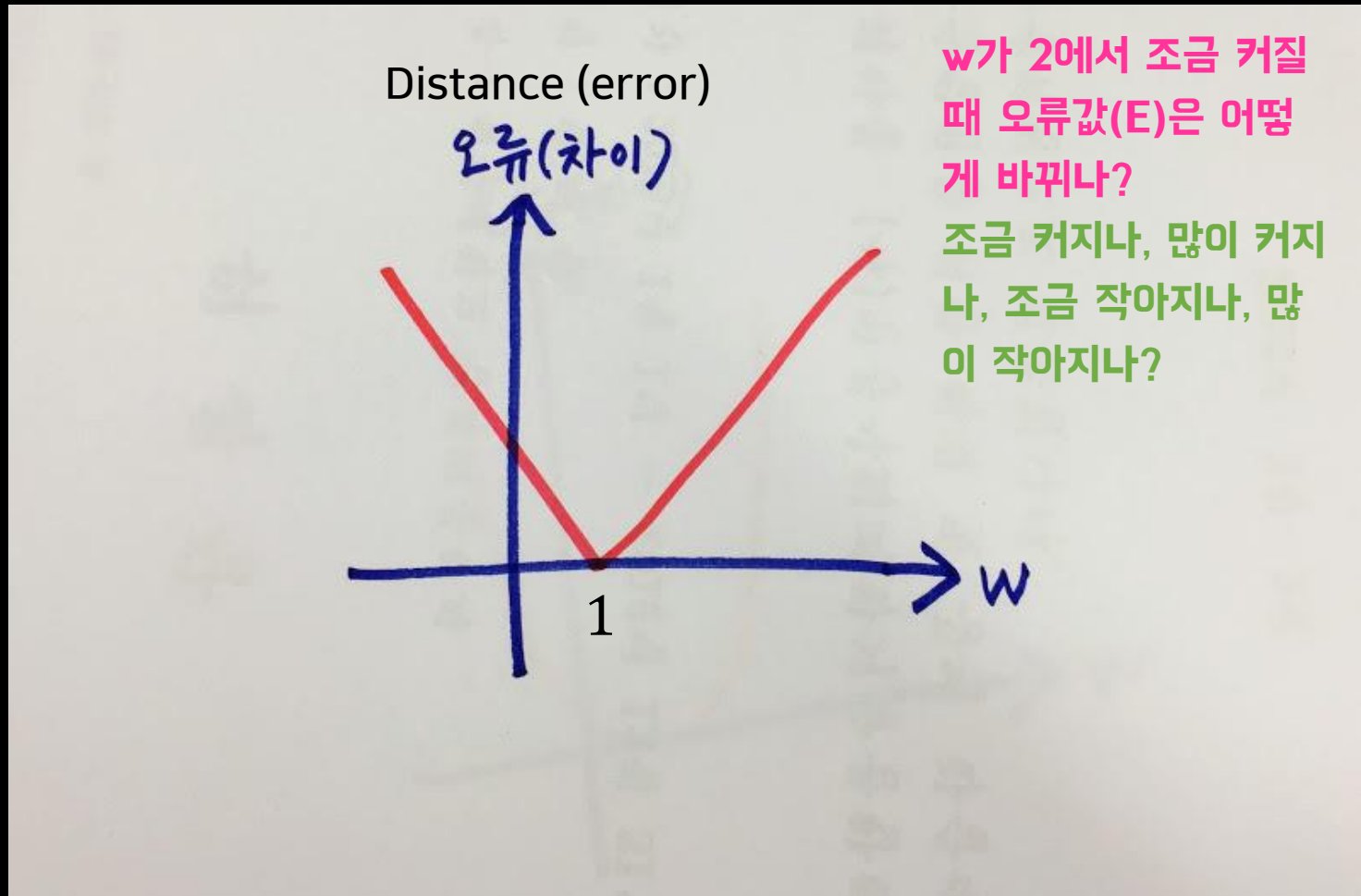
만일, 데이터가 3개라면

$$E = \frac{1}{3} \sum_{i=1}^3 |w x_i - y_i|$$

모두 더해서 평균

오류 그래프

$$E = \frac{1}{3} \sum_{i=1}^3 |w x_i - y_i|$$

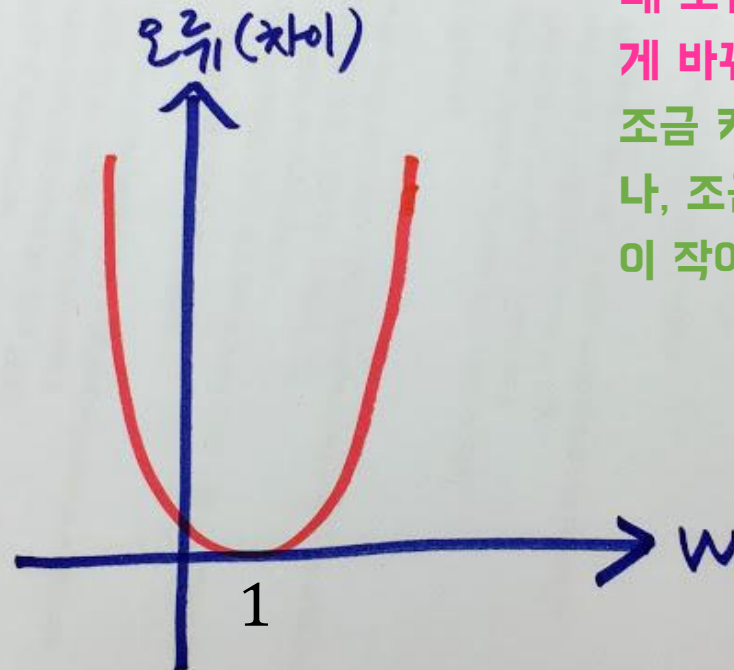


오류 그래프

$$E = \frac{1}{3} \sum_{i=1}^3 (w x_i - y_i)^2$$

MSE

Distance (error)



w가 2에서 조금 커질
때 오류값(E)은 어떻
게 바뀌나?

조금 커지나, 많이 커지
나, 조금 작아지나, 많
이 작아지나?

데이터가 m개일 경우

Mean
Square
Error

Prediction by computer

Correct answer

$$E = \frac{1}{m} \sum_{i=1}^m (\mathbf{w}x_i - y_i)^2$$

Get \mathbf{w} (and \mathbf{b}) to minimize the error (cost, loss) function.

In this example, $\mathbf{w}=1$

then our machine predicts correctly.

오류 그래프 해석

- w 가 변하면 오류 E 도 변한다.
- 오류를 줄이고 싶으면 w 를 적절히 바꾸면 된다.
- 현재 w 값에 따라 w 가 조금만 변해도 오류가 많이 변하는 곳도 있고,
- 어떤 곳에서는 w 가 변해도 오류가 아주 조금만 변하기도 한다.

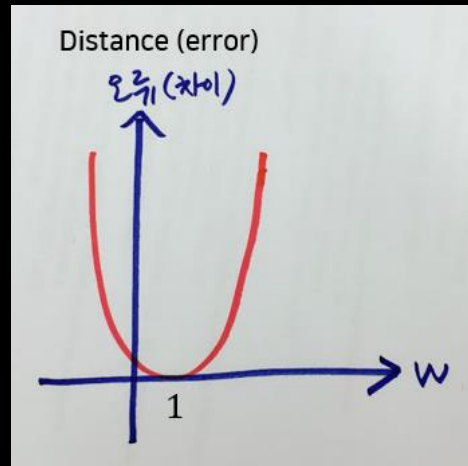
(Q) w 예상하기

- 현재 w 가 어떤 값일 때 w 를 조금 증가시켰더니 오류 E 가 아주 급격하게 늘어났다. 현재 w 값은?
- 현재 w 에서 w 를 조금 증가시켰더니 오류는 반대로 급격하게 감소하였다. 어디?
- w 를 변경해 보았지만 오류는 거의 변하지 않았다. 어디?

현재 w 가 어떤 값일 때

w 변화가 오류 E 에 미치는 영향

예를 들어, 현재 w 가 4인 곳에서는 w 가 조금만 늘려도
오류(E)는 아주 크게 늘어난다.

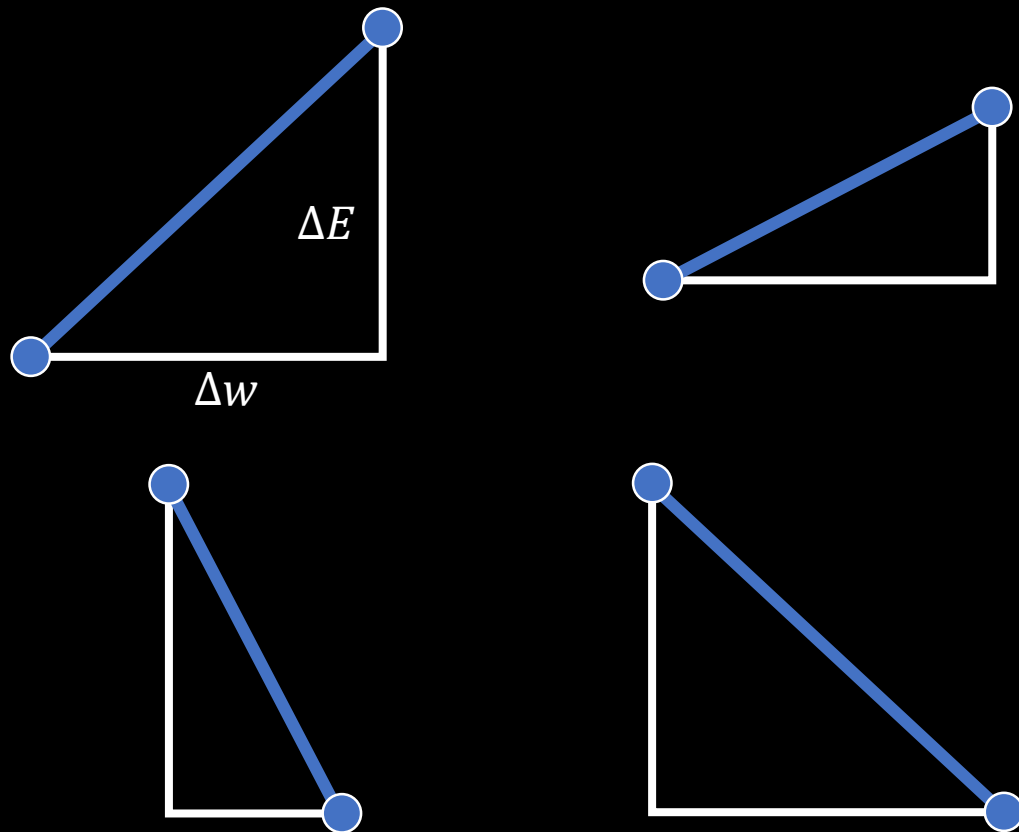


현재 w 가 어떤 값일 때

w 변화가 오류 E 에 미치는 영향,
기울기로 표현된다.

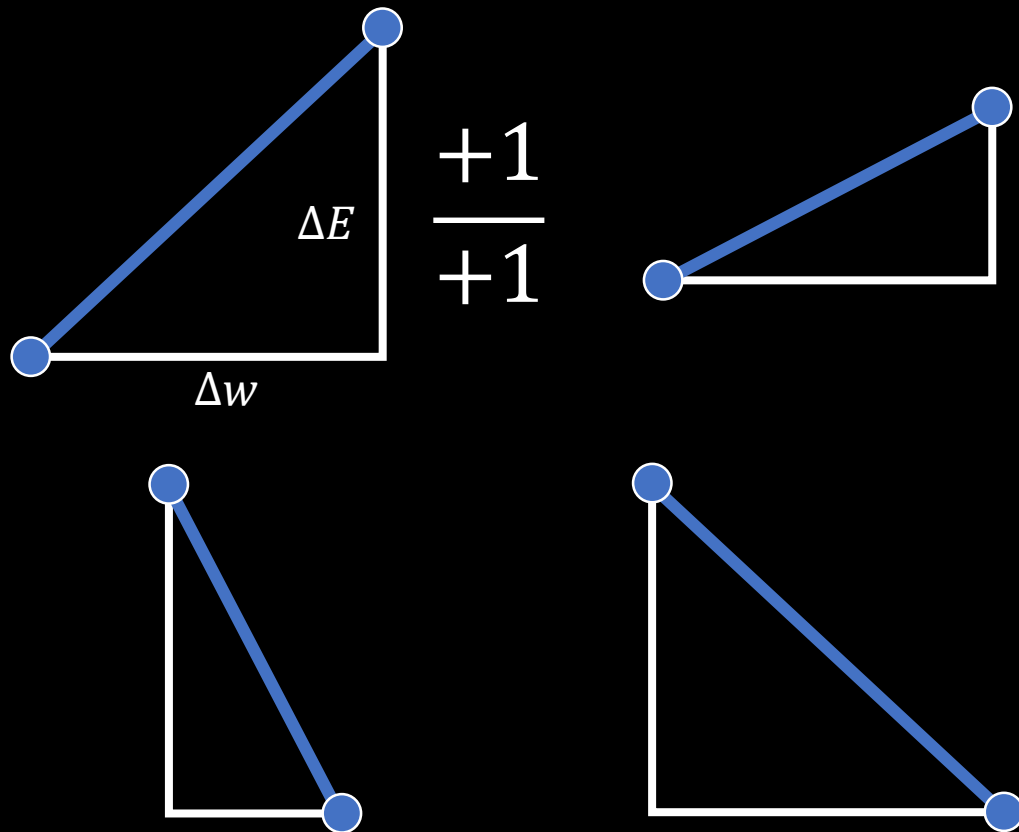
예) w 가 1만큼 늘어날 때 오류(E)는 2만큼 늘어날 경우 기울기는 1

기울기 의미(w 와 E 관계)



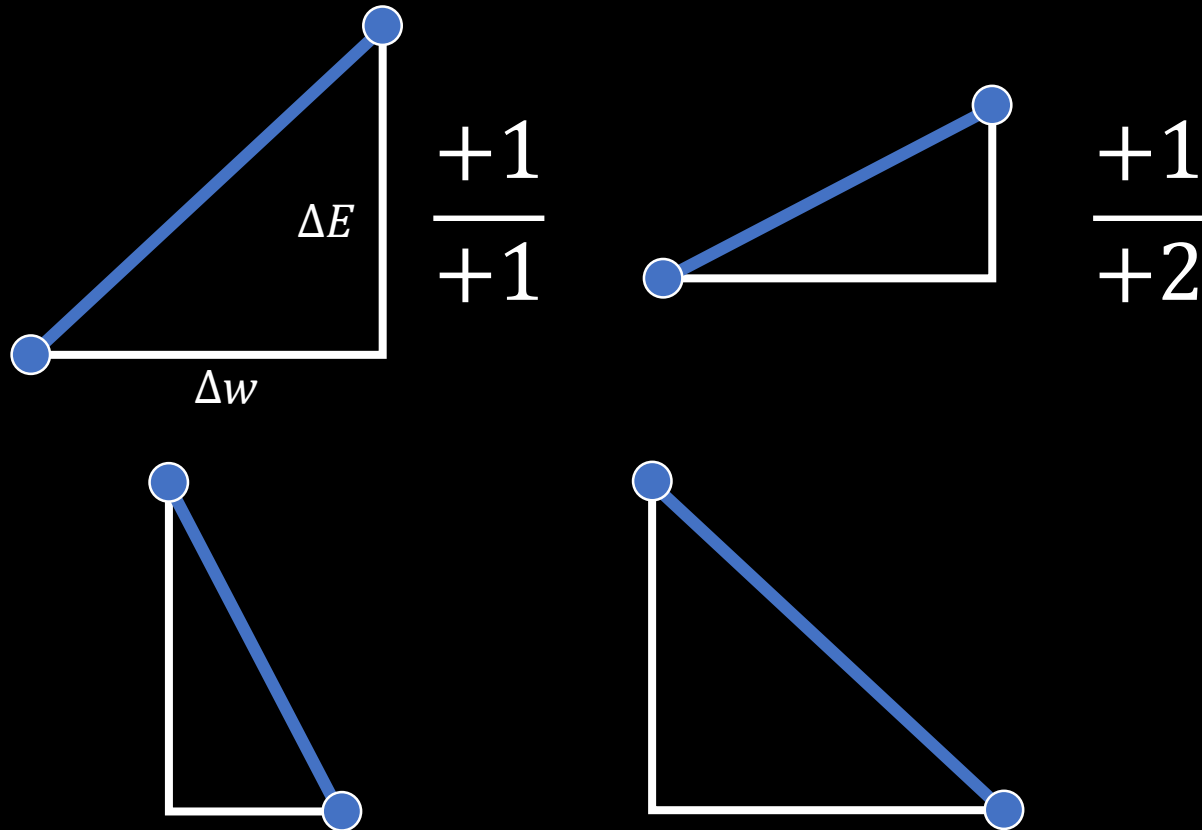
오류 그래프에서 두 점을 연결했을 때의 기울기

기울기 의미(w 와 E 관계)



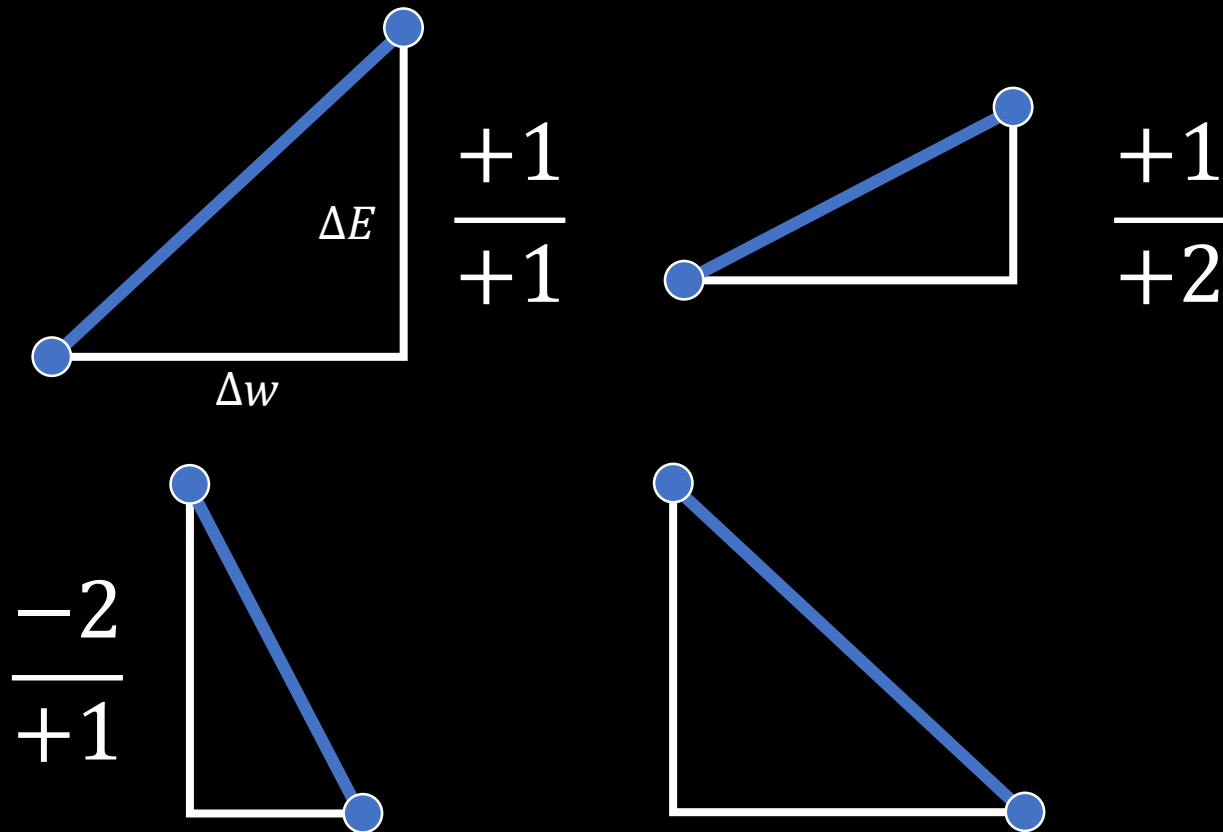
오류 그래프에서 두 점을 연결했을 때의 기울기

기울기 의미(w 와 E 관계)



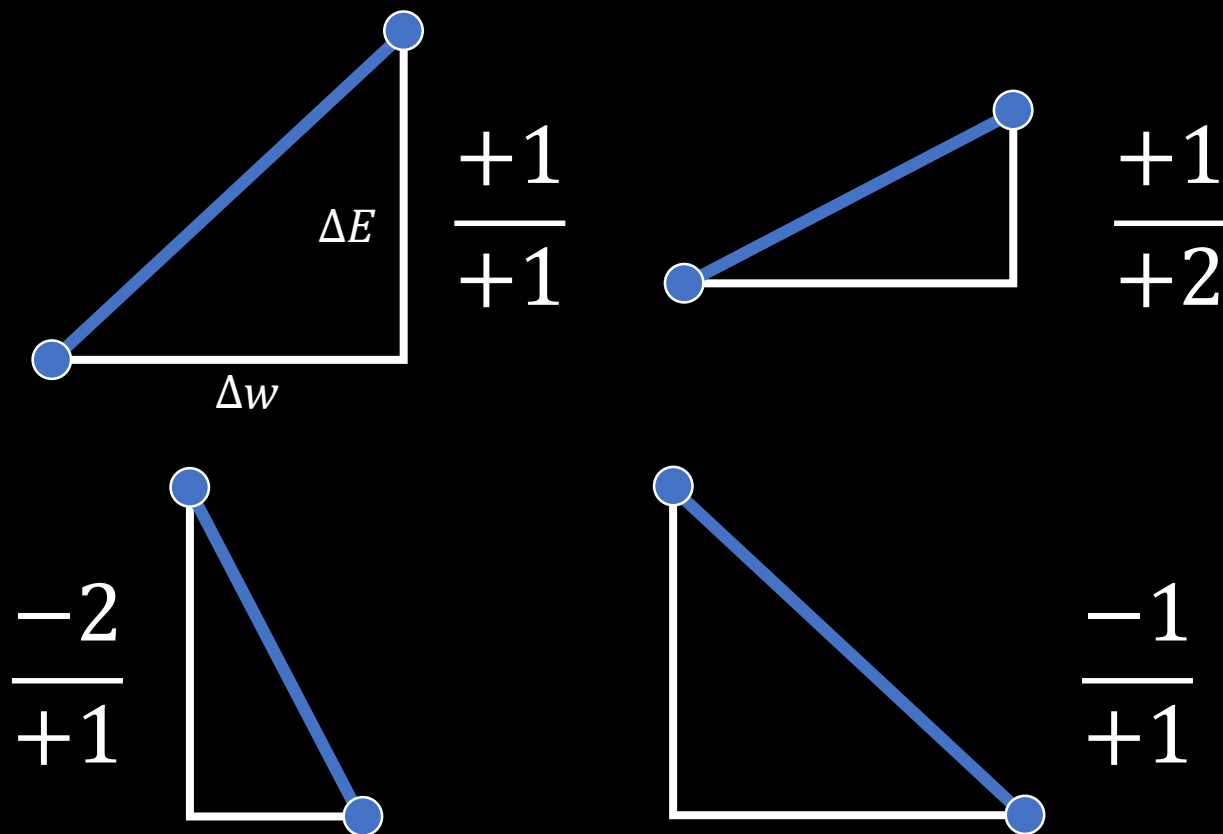
오류 그래프에서 두 점을 연결했을 때의 기울기

기울기 의미(w 와 E 관계)



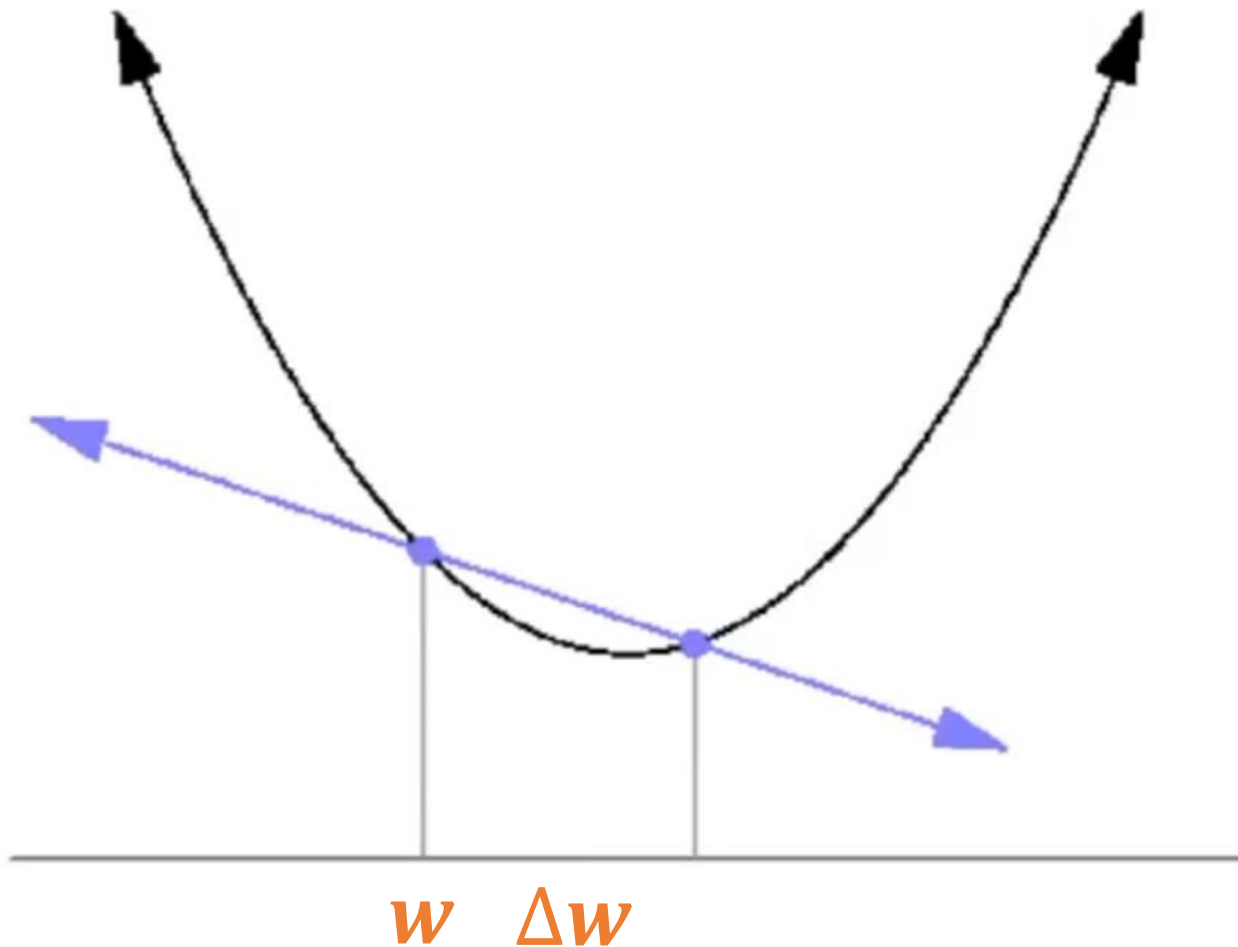
오류 그래프에서 두 점을 연결했을 때의 기울기

기울기 의미(w 와 E 관계)

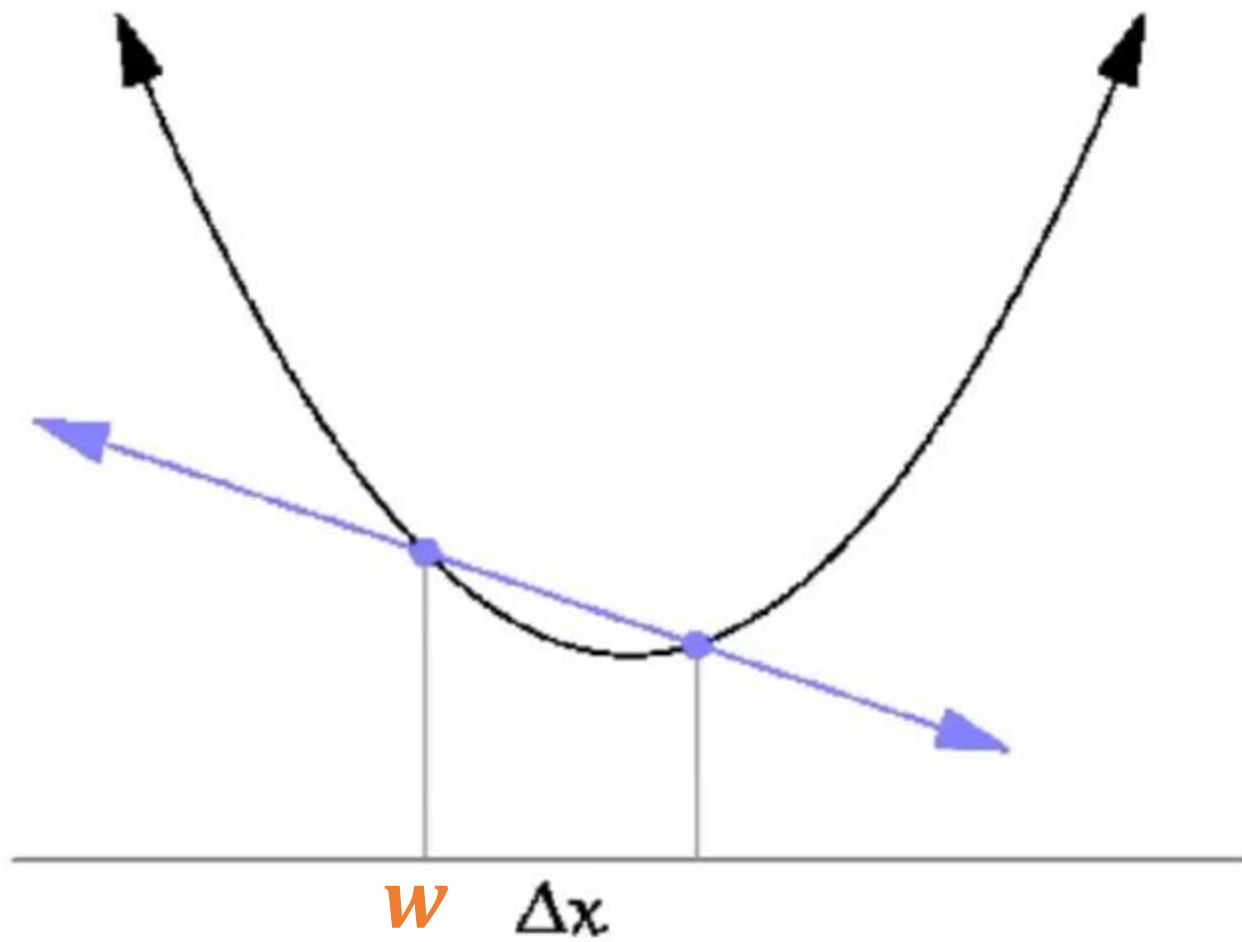


오류 그래프에서 두 점을 연결했을 때의 기울기

오류 그래프에서 기울기



오류 그래프에서 두 점을 연결했을 때의 기울기



Numerical differentiation

오류 그래프 모양을 몰라도,
공식 E 만 알면 기울기를
계산할 수 있다.
즉, w 변화가 E 에 미치는
영향을 알 수 있다.

(Q) 미치는 영향 구하기

$$E = (wx - y)^2$$

데이터 (x, y) 가 $(1, 1)$ 일 때
 $w=3$ 인 지점에서 w 변화가 E 에 미치는 영향, 즉, 기울기를 구하라.

(Q) 미치는 영향 구하기

$$E = (w1 - 1)^2$$

w : 3 \rightarrow E : 4

w : 3.00001 \rightarrow E : 4.00004

w 가 0.00001 증가할 때 E 는 0.00004 증가

w 가 1증가할 때 E 는 4 증가함을 의미

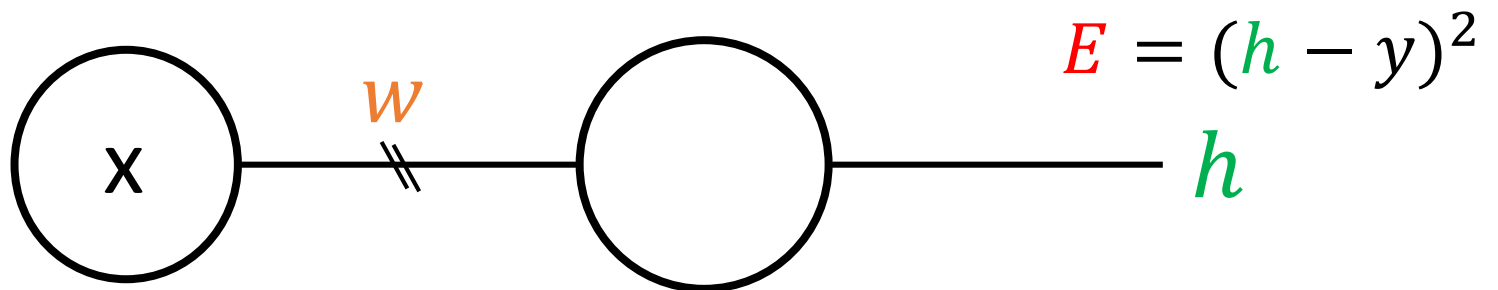
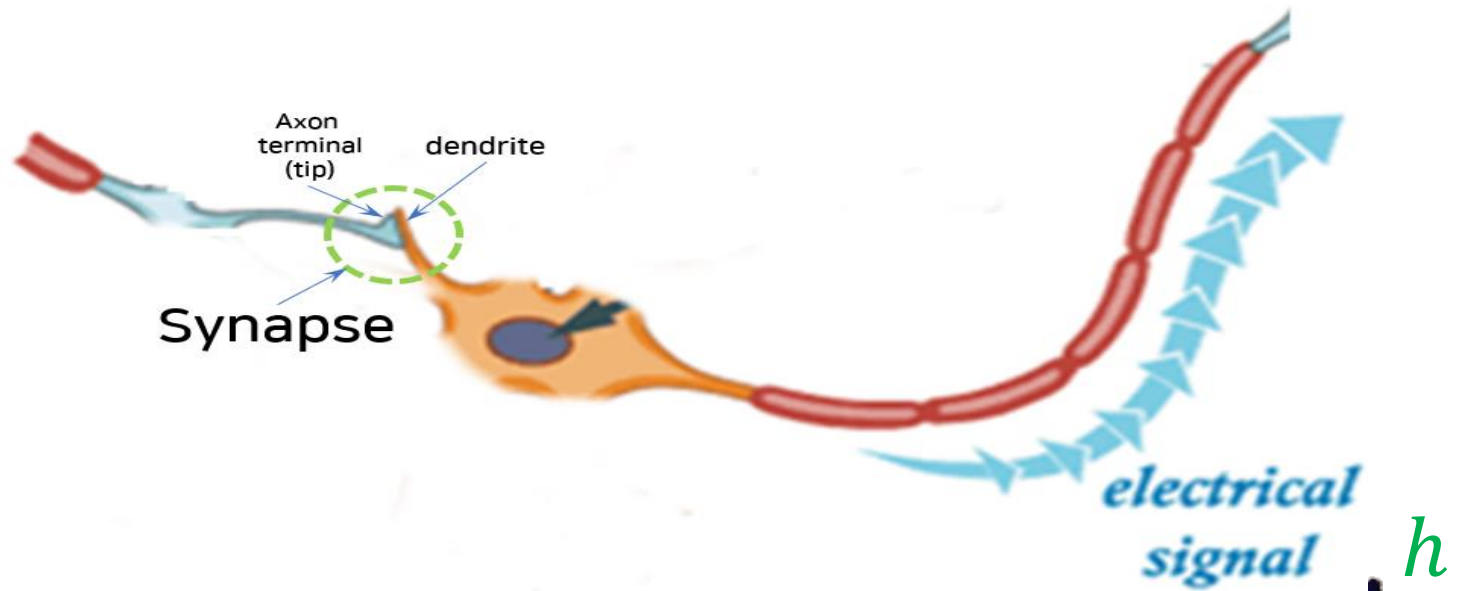
따라서 미치는 영향(기울기)=4

기울기= \square 치는 영향,
공식이 있을 때 어떻게 구한다?

기울기에 대한 생각

- 기울기가 4 → w 가 1만큼 증가하면 오류 E 는 4만큼 증가
- 따라서 오류를 줄이고 싶으면? w 를 감소시켜야 함을 의미
- 기울기가 -5 → w 가 1만큼 증가하면 오류는 5만큼 감소 뜻
- 따라서 오류를 줄이고 싶으면? w 를 증가시켜야 함을 의미

w 를 변화시키면서(튜닝)
오류 E 를 계속해서 줄이자.
(학습, Learning)



Learning is ...

- parameter(w) tuning to minimize the error/cost/loss function (**learning**)
- then machine can **predict** correct answers.

어떻게 '자동으로'

- 오류 E 를 최소화하는 w 값을 찾을까?

요약

- 리그레션(회귀)을 이해한다.
- 가설과 오류 그래프를 이해한다.
- 오류 그래프를 해석할 수 있다.
- 가중치가 오류에 미치는 영향을 안다.
- 기울기의 의미를 안다.