1. 머신러닝이란 무엇인가?

* 많은 rule을 가지고 있는 프로그래밍을 하기 위해 개발
* Data를 통해 학습을 해서 자동적으로 rule를 짤 수 있는 프로그램

1. Supervised learning

지도 학습 : label이 있는 상태에서 학습을 하는 머신러닝

Image labeling, email spam filter, predicting exam score

범위가 넓은 시험 점수 예측 같은 경우 🡪 regression

Pass/ non-pass 🡪 Classification

Grade (A,B,C,D) 🡪 Multi-label classification

연속이나 discrete이냐에 따라 regression 인지 classification인지 나뉘는 것 같음?

1. Unsupervised learning

Un-labeled data (data를 보고 스스로 학습 ) :clustering 이 예시

1. Linear Regression

Cost = (H(x) – y)^2

* 부호를 일정하게 양수로 표현가능 함
* 차이가 클 때 cost 가 커지는 가중치를 줄 수 있음

H(x) 🡪 이론값 ( 가설값)

m= 데이터의 개수

Cost\_function = 1/m \* sigma i to m (cost)

* 데이터들의 오차를 모두 더해주고 학습데이터의 숫자로 나누어 정규화!

**정규화는 왜 중요한가? (Q&A)**

Goal 🡪 Minimize cost

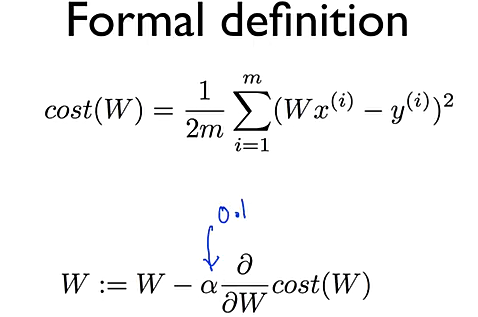
Minimize cost(W,b)

**Linear 에서는 W가 하나의 scalar 값 but curve fitting 할 때는 W가 하나의 1-d vector로 들어올 것 W = [w1,w2,w3,w4 …. ] (둘 다 일경우 행렬식으로 들어올수도 있음 🡪 수학적으로 유도해보기)**

**또는 x값 자체의 원소가 vector일 경우도 W의 shape가 변환될 것**

Gradient descent algorithm 이 보통 ML에서 많이 쓰임

* 궁금한점 다른 방법은 없는가?



Cost function 이 Convex function이여야지 local optimization 이 아닌 global optimization 이 됨

🡪convex function인지 먼저 확인을 해줘야함

1. Multi linear regression

1)cost 변환

Cost = w1x1+w2x2+w3x3 + …. w\_n\*x\_n

Cost\_function = 1/m \* sigma 1 to m (cost)

Matrix 형태로 식이 변환

X = [x1,x2,x3 … x\_n]

W = [w1,w2,w3, … w\_n]\_Transpose

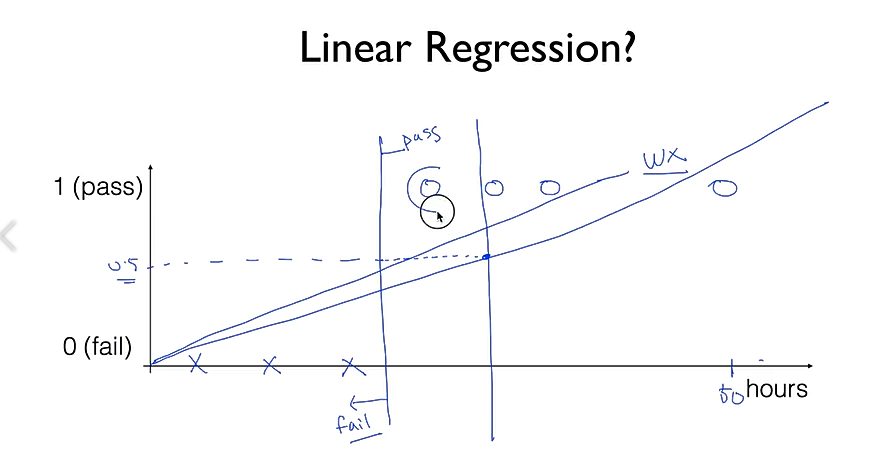
B=[b1,b2,b3 … b\_n]

H(X) = XW +B

출력이 여러 개로 되길 원하면 matrix 의 shape를 조절하면 댐 🡪 BPNN 이랑 같음

1. Logistic Regression (classification)

Binary classification 🡪 두개의 class를 분류하는 것 (흑/백)

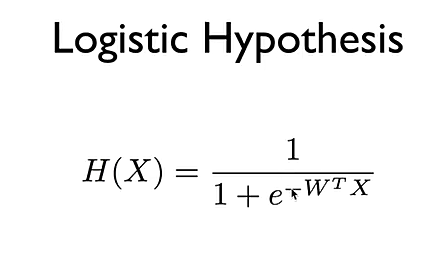


Linear regression 으로 직선을 만들어 threshold를 정할 시 위와 같이 예외상황이 생김

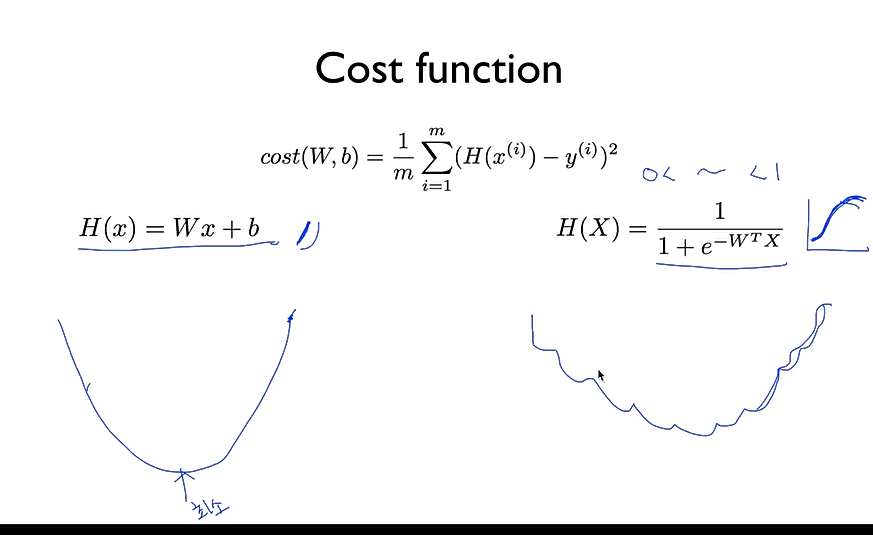
숫자가 커질 경우 문제가 생김 🡪 0~1 사이로 압축을 시켜주는 함수를 하나 고안

Logistic Hypothesis

* Sigmoid 고안



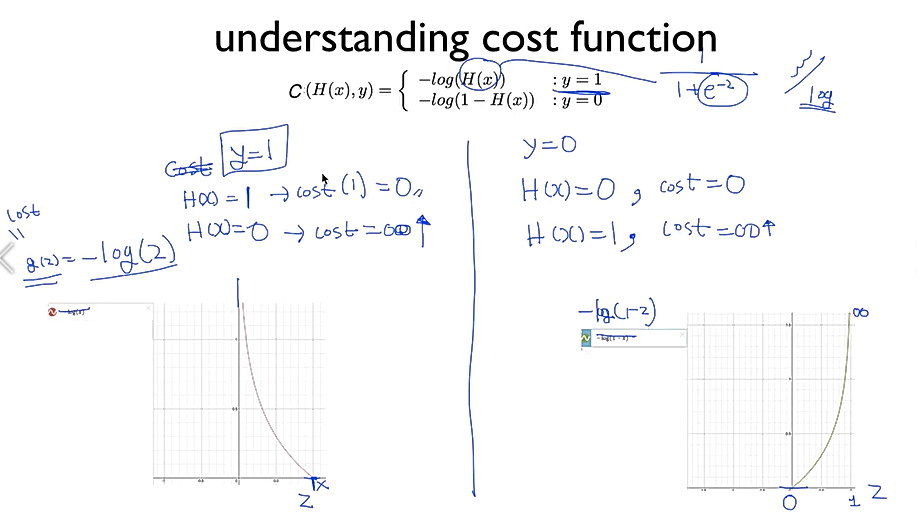
sigmoid 함수를 이용하여 함수의 값이 0~1사이에 존재하게 함



H(x) 가 sigmoid가 합성되었기 때문에 local minimum이 많이 생김

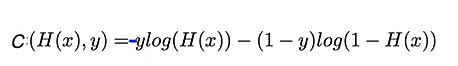
따라서 H(x)가 저런 모양이 되면 바로 gradient descendent 방법을 쓸 수 없음 (너무 많은 local min)

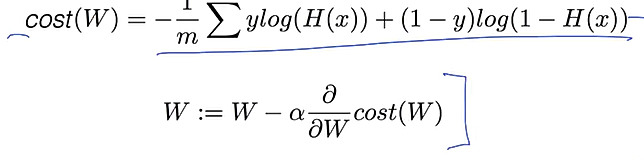
New cost function 고안 🡪 logistic function



실제 target ( y) 의 값이 1일 때 H(x) 가 1이면 cost 가 0, H(x) 가 0이면 cost가 무한으로 틀리게 될 경우 cost값이 점점 높아지기 때문에 적절한 cost function이라 할 수 있음

But 위와 같은 함수를 사용할 경우 조건이 많이 생겨 불편 따라서 다음과 같은 식을 고안

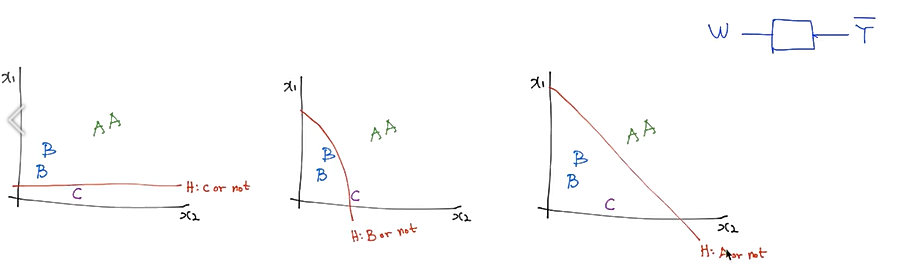




* 최종적으로 Class 를 구별하는 문제에서는 Linear regression 의 경우 적합하지 않기 때문에 logistic regression으로 수행 (0~1사이 값으로)
* 즉 Classification 의 경우 logistic regression 으로 분류! (Binary classification)

1. Softmax regression == Multinomial classification

여러 개의 Binary classification을 통해 multinomial도 구현할 수 있음



1. C or not 2) B or not 3) A or not

여러 개의 regression을 동시에 해야함 == matrix 형태의 연산